

si se desea, una secuencia de ácido nucleico que codifica para un péptido susceptible de ser utilizado con fines de aislamiento o purificación del péptido o proteína de fusión. Por tanto, en una realización particular, la construcción de ADN de la invención contiene, si se desea, una quinta secuencia de ácido nucleico que codifica para un péptido susceptible de ser utilizado con fines de aislamiento o purificación.

Prácticamente cualquier péptido o secuencia peptídica que permita el aislamiento o purificación del péptido o proteína de fusión puede ser utilizada, por ejemplo, una secuencia de polihistidina, o una secuencia peptídica reconocida por un anticuerpo monoclonal y que puede servir para purificar la proteína de fusión resultante por cromatografía de inmutafinidad, por ejemplo, péptidos etiqueta tales como c-myc, HA, E, FLAG, etc. [Using Antibodies: A laboratory manual. Ed Harlow and David Lane (1999). Cold Spring Harbor Laboratory Press. New Cork. Capítulo: Tagging proteins. pp. 347-377] y, en general, cualquier otra secuencia reconocida por un anticuerpo.

Dicha quinta secuencia de ácido nucleico puede estar situada en cualquier posición de la construcción de ADN de la invención excepto en la región correspondiente al extremo C-terminal de HlyA ya que, en ese caso, rompería la señal de secreción. A modo ilustrativo, dicha quinta secuencia de ácido nucleico podría estar situada entre dichas segunda y tercera secuencias de ácido nucleico, en donde el extremo 5' de dicha quinta secuencia de ácido nucleico está unido al extremo 3' de dicha segunda secuencia de ácido nucleico y el extremo 3' de dicha quinta secuencia de ácido nucleico está unido al extremo 5' de dicha tercera secuencia de ácido nucleico. Alternativamente, dicha quinta secuencia de ácido nucleico podría estar situada en la región correspondiente al extremo N-terminal de la proteína de fusión resultante o entre el producto de interés y el dominio de dimerización.

Para facilitar el reconocimiento de la proteína o péptido de fusión obtenido, la construcción de ADN de la invención también puede contener, si se desea, una sexta secuencia de ácido nucleico que codifica para un péptido susceptible de ser utilizado con fines de reconocimiento.

Prácticamente cualquier péptido o secuencia peptídica que permita el reconocimiento del péptido o proteína de fusión puede ser utilizada, por ejemplo, una secuencia peptídica reconocida por un anticuerpo monoclonal y que puede servir para reconocer la proteína de fusión resultante por técnicas de inmunodetección, por ejemplo,

péptidos etiqueta tales como c-myc, HA, E, FLAG y, en general, cualquier otra secuencia reconocida por un anticuerpo.

Dicha sexta secuencia de ácido nucleico puede estar situada en cualquier posición de la construcción de ADN de la invención excepto en la región correspondiente al extremo C-terminal de HlyA para evitar que se rompa la señal de secreción. A modo
5 ilustrativo, dicha sexta secuencia de ácido nucleico podría estar situada entre dichas segunda y tercera secuencias de ácido nucleico, en donde el extremo 5' de dicha sexta secuencia de ácido nucleico está unido al extremo 3' de dicha segunda secuencia de ácido nucleico y el extremo 3' de dicha sexta secuencia de ácido nucleico está unido al extremo
10 5' de dicha tercera secuencia de ácido nucleico. Alternativamente, dicha sexta secuencia de ácido nucleico podría estar situada en la región correspondiente al extremo N-terminal de la proteína de fusión resultante o entre el producto de interés y el dominio de dimerización.

Dichas quinta y sexta secuencias de ácido nucleico pueden estar separadas entre sí.
15 Alternativamente, en una realización particular, dichas quinta y sexta secuencias de ácido nucleico pueden estar unidas entre sí. En este caso, a modo ilustrativo, dicha sexta secuencia de ácido nucleico que codifica para un péptido susceptible de ser utilizado con fines de reconocimiento puede estar situada entre dichas tercera y quinta secuencias de ácido nucleico, en donde el extremo 5' de dicha sexta secuencia de ácido nucleico está
20 unido al extremo 3' de dicha quinta secuencia de ácido nucleico y el extremo 3' de dicha sexta secuencia de ácido nucleico está unido al extremo 5' de dicha tercera secuencia de ácido nucleico. Alternativamente, dichas secuencias pueden estar unidas entre sí en el orden inverso, en cuyo caso, dicha sexta secuencia de ácido nucleico que codifica para un péptido susceptible de ser utilizado con fines de reconocimiento está situada entre dichas
25 segunda y quinta secuencias de ácido nucleico, en donde el extremo 3' de dicha sexta secuencia de ácido nucleico está unido al extremo 5' de dicha quinta secuencia de ácido nucleico y el extremo 5' de dicha sexta secuencia de ácido nucleico está unido al extremo 3' de dicha segunda secuencia de ácido nucleico.

Si se desea, la construcción de ADN de la invención puede contener, además, una
30 secuencia de nucleótidos que codifica para una secuencia de aminoácidos susceptible de ser escindida específicamente por medios enzimáticos o químicos con el fin de liberar la proteína dimerica de interés una vez aislada la proteína de fusión. En este caso, la construcción de ADN de la invención puede incluir, además, una séptima secuencia de

ácido nucleico que comprende una secuencia de nucleótidos que codifica para una secuencia de aminoácidos susceptible de ser escindida específicamente por medios enzimáticos o químicos. Prácticamente cualquier secuencia de aminoácidos susceptible de ser escindida específicamente por medios enzimático o químicos puede ser utilizada. En una realización particular, dicha séptima secuencia de ácido nucleico comprende una secuencia de nucleótidos que codifica para un sitio de reconocimiento de una proteasa, por ejemplo, una enteroquinasa, Arg-C endoproteasa, Glu-C endoproteasa, Lys-C endoproteasa, Factor de coagulación Xa y similares. En otra realización particular, dicha séptima secuencia de ácido nucleico comprende una secuencia de nucleótidos que codifica para un sitio susceptible de ser específicamente escindido por un reactivo químico tal como, por ejemplo, bromuro de cianógeno que escinde restos de metionina o cualquier otro reactivo químico apropiado.

Dicha séptima secuencia de ácido nucleico se encuentra, generalmente, a continuación del extremo 3' de dicha segunda secuencia de ácido nucleico, en cualquier posición entre las secuencias segunda y tercera de ácido nucleico, de manera que por medios enzimáticos o químicos se pueda escindir la proteína dímera de interés.

La construcción de ADN de la invención puede obtenerse mediante el empleo de técnicas ampliamente conocidas en el estado de la técnica [Sambrook et al., "Molecular cloning, a Laboratory Manual", 2nd ed., Cold Spring Harbor Laboratory Press, N.Y., 1989 Vol 1-3]. Dicha construcción de ADN de la invención puede incorporar, operativamente unida, una secuencia reguladora de la expresión, constituyendo de este modo un cassette de expresión.

Por tanto, en otro aspecto, la invención proporciona al menos un cassette de expresión que comprende al menos una construcción de ADN de la invención operativamente unida a una secuencia de control de expresión. Las secuencias de control son secuencias que controlan y regulan la transcripción y, en su caso, la traducción del producto de interés, e incluyen secuencias promotoras (*pT7*, *plac*, *pBAD*, *ptet*, etc.), secuencias codificantes para reguladores transcripcionales (*lacI*, *tetR*, *araC*, etc.), secuencias de unión a ribosomas (RBS), y/o secuencias terminadoras de transcripción (*t1t2*, etc.), etc. En una realización particular, dicha secuencia de control de expresión es funcional en bacterias, en particular, en bacterias Gram negativas.

La construcción de ADN de la invención, o el cassette de expresión proporcionado por esta invención, pueden ser insertados en un vector apropiado. Por

tanto, en otro aspecto, la invención se relaciona con un vector, tal como un vector de expresión, que comprende al menos una construcción de ADN o al menos un cassette de expresión. La elección del vector dependerá de la célula hospedadora en la que se va a introducir posteriormente. A modo de ejemplo, el vector donde se introduce dicha
5 secuencia de ADN puede ser un plásmido o un vector que, cuando se introduce en una célula hospedadora, se integra o no en el genoma de dicha célula. La obtención de dicho vector puede realizarse por métodos convencionales conocidos por los técnicos en la materia [Sambrook et al., 1989, citado *supra*].

Ventajosamente, dicho vector comprende, además, un marcador o gen que
10 codifica para un motivo o para un fenotipo que permita la selección de la célula hospedadora transformada con dicho cassette de expresión. Ejemplos ilustrativos de dichos marcadores que podrían estar presentes en el cassette de expresión de la invención incluyen genes de resistencia a antibióticos, por ejemplo, ampicilina, tetraciclina, kanamicina, cloranfenicol, espectinomycin, etc., o genes de resistencia a
15 compuestos tóxicos (telurito, mercurio, etc.).

En otro aspecto, la invención se relaciona con una bacteria, en particular, una bacteria Gram negativa, que comprende al menos una construcción de ADN de la invención o al menos un cassette de expresión de la invención, o al menos un vector de la invención, en adelante bacteria de la invención. Dicha bacteria debe tener el sistema
20 exportador de hemolisina (Hly) de *E. coli* para lo cual, si no lo tiene de forma nativa, se debe proporcionar dicho sistema a la bacteria transformándola con un vector que contenga los genes HlyB y HlyD, por ejemplo, el plásmido pVDL9.3 [Fernández, L.A. et al., Applied and Environmental Microbiology, Nov. 2000, 5024-5029]. Prácticamente cualquier bacteria Gram negativa, por ejemplo, *E. coli*, *Salmonella*
25 *tiphymurium*, *Pseudomonas aeruginosa*, *Pseudomonas putida*, etc., puede ser transformada con la construcción de ADN de la invención o con el cassette de expresión de la invención. Para ello, las señales promotoras, reguladoras, marcadoras y orígenes de replicación deben ser optimizados para cada especie bacteriana. En una realización particular, dicha bacteria Gram negativa es *Escherichia coli*.

30 La construcción de ADN de la invención puede ser utilizada para producir productos de interés. Por tanto, en otro aspecto, la invención se relaciona con un método para producir un producto de interés, en forma de una proteína de fusión dimérica, que comprende crecer una bacteria de la invención bajo condiciones que permiten la

producción y excreción al medio de cultivo de dicho producto de interés en forma de una proteína de fusión dimérica. En una realización particular de la invención dicha proteína de fusión dimérica comprende dos productos de interés. Por tanto, en un aspecto aún más preferido de la invención se obtendría una proteína de fusión dimérica por expresión de las secuencias de ácido nucleico contenidas en al menos una construcción de ADN de la invención, o al menos un cassette de expresión de la invención, o al menos un vector de la invención. Las condiciones para optimizar el cultivo de la bacteria de la invención dependerán de la bacteria utilizada.

Si se desea, el método para producir un producto de interés proporcionado por esta invención incluye, además, el aislamiento y purificación de dicha proteína de fusión dimérica. En este caso, la construcción de ADN de la invención incluye, además, dicha séptima secuencia de ácido nucleico previamente definida que comprende una secuencia de nucleótidos que codifica para una secuencia de aminoácidos susceptible de ser escindida específicamente por medios enzimáticos o químicos con el fin de liberar el producto de interés. En una realización particular, dicha secuencia de nucleótidos codifica para un sitio de reconocimiento de una proteasa, por ejemplo, una enteroquinasa, Arg-C endoproteasa, Glu-C endoproteasa, Lys-C endoproteasa, Factor de coagulación Xa y similares. En otra realización particular, dicha secuencia de nucleótidos codifica para un sitio susceptible de ser específicamente escindido por un reactivo químico tal como, por ejemplo, bromuro de cianógeno que escinde restos de metionina o cualquier otro reactivo químico apropiado.

En otro aspecto, la invención se relaciona con una proteína de fusión dimérica obtenible por expresión de al menos una secuencia de ácido nucleico contenida en al menos una construcción de ADN de la invención, en la que cada monómero comprende:

- (i) la secuencia de aminoácidos de un producto de interés;
- (ii) una secuencia de aminoácidos correspondiente a un dominio de dimerización; y
- (iii) la secuencia de aminoácidos de la HlyA de *E. coli* o de un fragmento de dicha proteína que comprende la señal de reconocimiento del mecanismo de secreción del sistema transportador de hemolisina (Hly) de *E. coli*.

De forma más concreta, cada monómero de la proteína de fusión dimérica de la invención comprende

- 5 (i) un producto de interés, por ejemplo, una enzima, un inhibidor enzimático, una hormona, una molécula implicada en la adhesión y/o señalización celular y compuesta por dominios, por ejemplo, una inmunoglobulina, un antígeno inmunogénico, tal como una proteína o un fragmento antigénico de la misma procedente de un patógeno, por ejemplo, de un patógeno viral, bacteriano, de un parásito, etc., que puede causar infecciones en seres humanos o animales, un agente terapéutico,
- 10 por ejemplo, un antígeno específico de tumor, un antígeno de una enfermedad auto-inmune, etc., o una molécula inmunoreguladora, por ejemplo, un factor de crecimiento, una citoquina, tal como una interleuquina, un interferón, etc.; en una realización particular, dicho producto de interés es un minianticuerpo susceptible de ser utilizado con fines terapéuticos, de diagnóstico o de investigación;
- 15 (ii) un dominio de dimerización, tal como una hélice peptídica, una estructura de doble arrollamiento helicoidal (coiled coil), o, en general, cualquier secuencia peptídica que promueva la dimerización en las proteínas que la contengan. En una realización particular, dicho dominio de dimerización comprende la cremallera de leucina del factor de transcripción GCN4 de levadura; y
- 20 (iii) la secuencia completa de aminoácidos de HlyA de *E. coli*, o alternativamente, un fragmento de HlyA de *E. coli* que comprende la señal de reconocimiento del mecanismo de secreción del sistema transportador de Hly de *E. coli*.
- 25

Cada monómero de la proteína de fusión dimérica de la invención también puede contener, si se desea, (a) un péptido espaciador entre el producto de interés y el dominio de dimerización; ventajosamente, dicho péptido espaciador es un péptido con

30 flexibilidad estructural, por ejemplo, un péptido que contiene repeticiones de restos de aminoácidos, tales como Gly-Gly-Gly-Ser, o cualquier otra repetición de restos de aminoácidos adecuada, o bien la región bisagra de un anticuerpo; en una realización particular, dicho péptido espaciador flexible comprende la región bisagra de un

anticuerpo; y/o (b) un péptido para facilitar el aislamiento o purificación del péptido o proteína de fusión, por ejemplo, una secuencia de polihistidina, o una secuencia peptídica reconocida por un anticuerpo monoclonal y que puede servir para purificar la proteína de fusión resultante por cromatografía de inmunoafinidad, por ejemplo, péptidos etiqueta tales como c-myc, HA, E, FLAG, y, en general, cualquier otra secuencia reconocida por un anticuerpo; y/o (c) un péptido que permita el reconocimiento del péptido o proteína de fusión, por ejemplo, una secuencia peptídica reconocida por un anticuerpo monoclonal y que puede servir para reconocer la proteína de fusión resultante por técnicas de inmunodetección, por ejemplo, péptidos etiqueta tales como c-myc, HA, E, FLAG y, en general, cualquier otra secuencia reconocida por un anticuerpo; y/o (d) una secuencia de aminoácidos susceptible de ser escindida específicamente por medios enzimáticos, por ejemplo, una secuencia de aminoácidos que constituye un sitio de reconocimiento de una proteasa, por ejemplo, una enteroquinasa, Arg-C endoproteasa, Glu-C endoproteasa, Lys-C endoproteasa, Factor de coagulación Xa y similares, o una secuencia de aminoácidos susceptible de ser específicamente escindida por un reactivo químico tal como, por ejemplo, bromuro de cianógeno y similares.

El sistema de producción de proteínas de fusión diméricas proporcionado por esta invención es particularmente útil para la producción de proteínas implicadas en una actividad de unión, por ejemplo, en las interacciones proteína-ADN o antígeno-anticuerpo, puesto que puede intensificar su avidez.

En otra realización aún más preferida de la invención, las construcciones de ADN de la invención sirven para la creación y expresión de una librería de proteínas diméricas, por ejemplo de minianticuerpos. Un ejemplo particular de esta realización sería utilizar los dímeros de minianticuerpos así producidos en procesos de selección de moléculas con capacidad de unión a un antígeno determinado.

En otra realización de la invención, el sistema de producción de proteínas diméricas sirve para la producción de heterodímeros entre dos moléculas con capacidad de unión para antígenos diferentes o epítomos diferentes del mismo antígeno, preferentemente dos minianticuerpos, o un minianticuerpo y otro tipo de molécula, como puede ser, pero no se limita a, una toxina, una droga antitumoral, un enzima o moléculas implicadas en marcaje o detección. Así, un ejemplo particular de esta realización sería utilizar estos dímeros para el marcaje de células tumorales o el

transporte de moléculas con actividad antitumoral al tumor. La producción de este tipo de heterodímeros presenta importantes aplicaciones en diagnóstico y terapia.

Una ventaja del sistema proporcionado por esta invención radica en que permite producir proteínas tóxicas para un huésped bacteriano. Como es conocido, la expresión por métodos recombinantes de proteínas tóxicas para un huésped bacteriano es muy complicada o prácticamente imposible cuando dicha proteína tóxica expresada no es exportada desde la bacteria hacia el exterior. Con el método proporcionado por esta invención se puede expresar una proteína tóxica o previamente inexpressable, o expresada a bajos niveles, para producir la proteína deseada en cantidades utilizables.

El siguiente ejemplo ilustra la invención sin que deba ser considerado como limitativo del alcance de la misma.

EJEMPLO 1

Producción de minianticuerpos diméricos con elevada afinidad secretados por el sistema de transporte de hemolisina (Hly) de *E. coli*

En este ejemplo se describe la secreción de minianticuerpos diméricos en sobrenadantes de cultivos de *E. coli* que emplean el sistema de transporte de hemolisina (Hly). En primer lugar, se demostró que la dimerización puede conseguirse por ingeniería genética en el sistema de transporte de la Hly. Para ello se insertó una hélice α anfipática (es decir, el dominio de cremallera de leucina del factor de transcripción de levaduras GCN4) en el extremo N terminal de una versión marcada (E-tag) del dominio C terminal de 23 kDa de la hemolisina (EHlyA). Se comprobó que el polipéptido resultante (ZEHlyA) se secretaba eficazmente por las células de *E. coli* y se acumulaba en el medio de cultivo como un dímero estable. Después se utilizaron los vectores derivados de 'EHlyA y 'ZEHlyA para la secreción de los dominios V_{HH} de inmunoglobulinas obtenidos de anticuerpos de camello. Se secretaron los híbridos V_{HH} -EHlyA y V_{HH} -ZEHlyA y se encontraron en el medio extracelular como monómeros y dímeros, respectivamente. Cuando se compararon con sus homólogos monoméricos, las moléculas diméricas V_{HH} -ZEHlyA mostraron propiedades superiores de unión a su antígeno relacionado, con un aumento de 10 veces en su afinidad funcional (avidez). Este procedimiento permite obtener fácilmente minianticuerpos V_{HH} monoméricos y diméricos con alta avidez a partir de los sobrenadantes de cultivos de *E. coli*, facilitando

así la selección y la purificación de alto rendimiento de clones de V_{HH} a partir de grandes librerías.

1. MATERIALES Y MÉTODOS

5 **Cepas bacterianas, crecimiento y condiciones de inducción.** Las cepas de *E. coli* K-12 empleadas eran DH5 α F' (*supE44* Δ (*lacZYA-argF*)U169 Φ 80(*lacZ* Δ M15) *hsdR17 recA1 endA1 gyrA96 thi1 relA1*; Invitrogen) para la clonación y la propagación de los plásmidos y HB2151 (Δ *lac-pro*, *ara*, *nal^r*, *thi*, F'*proAB lacI^q lacZ* Δ M15) [Carter, P., H. Bedouelle, and G. Winter 1985. Improved oligonucleotide site-directed mutagenesis using M13 vectors. *Nucleic Acids Res.* 13:4431-4443] para la expresión de proteínas. Las bacterias que contenían los plásmidos indicados en cada caso se hicieron crecer a 30°C en placas de agar LB [Miller, J. H. 1992. A short course in bacterial genetics: a laboratory manual and handbook for *Escherichia coli* and related bacteria. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York] que contenía glucosa al 2% (p/v) (para reprimir el promotor *lac*) y los antibióticos apropiados para la selección de plásmidos. Para la inducción de los híbridos HlyA, se inocularon colonias individuales en el medio LB líquido que contenía glucosa al 2% (p/v) y se hicieron crecer a 30°C o a 37°C hasta que a la densidad óptica a 600 nm ($DO_{600\text{ nm}}$) alcanzó un valor de 0,5 aproximadamente. En ese punto, las bacterias se recogieron por centrifugación, se resuspendieron a la misma densidad en LB que contenía isopropil-1-tio- β -D-galactósido (IPTG) 0,3 mM y se incubaron (a 30°C o a 37°C) con agitación (160 r.p.m.) durante un periodo de tiempo comprendido entre 4 y 16 h. Se recogieron los sobrenadantes del cultivo tras la eliminación de las células de *E. coli* por centrifugación (10.000xg, 10 min) y se añadió 1/10 del volumen de tampón fosfato salino (PBS) 10X concentrado [PBS: 25 Na_2HPO_4 8 mM, KH_2PO_4 1,5 mM, KCl 3 mM, NaCl 137 mM, pH 7,0]. Los sobrenadantes del cultivo se emplearon directamente para realizar inmunoensayos o se almacenaron a -80°C hasta su utilización. Los antibióticos añadidos al medio de cultivo para la selección de plásmidos fueron la ampicilina (Ap; 150 μ g/ml) y el cloranfenicol (Cm; 30 μ g/ml).

30

Plásmidos y oligonucleótidos. Se emplearon métodos estándar para la manipulación y el aislamiento del ADN, la amplificación por PCR y la secuenciación del ADN [Ausubel,

F. M., R. Brent, R. E. Kingston, D. D. Moore, J. G. Seidman, J. A. Smith, and K. Struhl 1994. Current Protocols in Molecular Biology. John Wiley & Sons, New York; Sambrook, J., E. Fritsch, and T. Maniatis 1989. Molecular cloning. A laboratory manual. Cold Spring Harbor Laboratory Press, New York]. Los oligonucleótidos se obtuvieron de Sigma Genosys (Reino Unido) o de Isogen Bioscience BV (Países Bajos). Los plásmidos pEHlyA (Ap^r), pEHlyA2-SD (Ap^r) y pVDL9.3 (Cm^r) ya han sido descritos [Fernández, L. A., and V. De Lorenzo 2001. Formation of disulphide bonds during secretion of proteins through the periplasmic-independent type I pathway Mol Microbiol. 40:332-46; Fernández, L. A., I. Sola, L. Enjuanes, and V. de Lorenzo 2000. Specific secretion of active single-chain Fv antibodies into the supernatants of *Escherichia coli* cultures by use of the hemolysin system Appl Environ Microbiol. 66:5024-5029; Tzschaschel, B. D., C. A. Guzman, K. N. Timmis, and V. de Lorenzo 1996. An *Escherichia coli* hemolysin transport system-based vector for the export of polypeptides: Export of Shiga-like toxin IIeB subunit by *Salmonella typhimurium aroA*. Nature Biotechnology. 14:765-769]. El plásmido pZEHlyA (Ap^r) se obtuvo insertando en el sitio único *SaII* de pEHlyA un fragmento de ADN de 170 pb que codifica para el dominio ZIP amplificado por PCR y digerido con *SaII*. El mapa del plásmido pZEHlyA se muestra en la Figura 6. La amplificación por PCR del dominio ZIP se llevó a cabo con la ADN polimerasa Vent (New England Biolabs), usando como molde 1 ng de pCLZIP (Codon Genetic Systems, GmbH), y como cebadores los oligonucleótidos identificados como SEQ ID NO: 5 y SEQ ID NO: 6, que incorporaban dos sitios *SaII* flanqueando el producto amplificado. El plásmido pZEHlyA2-SD (Ap^r) se obtuvo insertando en el sitio único *SaII* de pEHlyA2-SD el fragmento de ADN de 170 pb que codifica para ZIP obtenido por digestión con *SaII* de pZEHlyA. El mapa del plásmido pZEHlyA2-SD se muestra en la Figura 7. Se seleccionó la orientación del fragmento de ADN ZIP que producía una inserción interna en el dominio C-HlyA marcado con E de pZEHlyA y pZEHlyA2-SD tras la secuenciación del ADN. Los fragmentos de ADN de aproximadamente 0,3 kb que codificaban para los dominios V_{HH}, V_{amy} y V_{tx} se amplificaron por PCR con la ADN polimerasa Vent, usando como molde 1 ng de los fagémidos A100R3A2 (anti- α -amilasa) o R3E5 (vacuna antitetánica), respectivamente, y como cebadores los oligonucleótidos identificados como SEQ ID NO: 7 y SEQ ID NO: 8.

Los productos amplificados de ADN que codifican para V_{amy} y V_{tx} contenían los sitios de restricción flanqueantes *NcoI* y *SfiI*, lo que permitía su clonación en los mismos sitios de pEHlyA2-SD y pZEHlyA2-SD, generando así pV_{amy}HlyA, pV_{tx}HlyA, pV_{amy}ZhlyA y pV_{tx}ZHlyA. Los mapas de los plásmidos pV_{amy}HlyA y pV_{amy}ZhlyA se muestran en las Figuras 8 y 9, respectivamente. Los fagémidos A100R3A2 y R3E5 fueron proporcionados por el Dr. Hennie Hoogenboon (Dyax Co., EE. UU.). Ambos fagémidos son derivados de pCANTAB6 (Cambridge Research Biochemicals) que contienen los dominios V_{HH} de camélidos clonados entre los sitios *SfiI* y *NotI*.

- 10 **Electroforesis e inmunotransferencia de proteínas.** Se llevó a cabo la electroforesis en geles de dodecil sulfato sódico - poliacrilamida (SDS-PAGE) usando geles de apilamiento al 4% y de separación al 10% (acrilamida:bisacrilamida 29:1; Bio-Rad), utilizando el sistema de electroforesis Miniprotean® (Bio-Rad) y siguiendo los protocolos estándar [Ausubel, F. M., R. Brent, R. E. Kingston, D. D. Moore, J. G.
- 15 Seidman, J. A. Smith, and K. Struhl 1994. Current Protocols in Molecular Biology. John Wiley & Sons, New York; Fraile, S., F. Roncal, L. A. Fernandez, and V. de Lorenzo 2001. Monitoring Intracellular Levels of XylR in *Pseudomonas putida* with a Single-Chain Antibody Specific for Aromatic-Responsive Enhancer-Binding Proteins J Bacteriol. 183:5571-9]. Para la inmunotransferencia, las proteínas se transfirieron a una
- 20 membrana de difluoruro de polivinilideno (Immobilon-P, Millipore) usando un aparato de electroforesis por transferencia semiseca (Bio-Rad). La membrana se bloqueó en tampón MTP (leche desnatada al 3% p/v, Tween 20 al 0,1% v/v en PBS) y se detectaron los polipéptidos marcados con E con anticuerpo monoclonal anti-E marcado con peroxidasa (0,2 µg/ml en tampón MTP; Amersham Bioscience). El conjugado
- 25 anticuerpo-POD unido se reveló mediante quimioluminiscencia, tal como ya se ha descrito [Fraile, S., F. Roncal, L. A. Fernandez, and V. de Lorenzo 2001. Monitoring Intracellular Levels of XylR in *Pseudomonas putida* with a Single-Chain Antibody Specific for Aromatic-Responsive Enhancer-Binding Proteins J Bacteriol. 183:5571-9]. La membrana se expuso a una película de rayos X (X-OMAT®, Kodak) o a
- 30 ChemiDoc® (Bio-Rad) para la cuantificación por quimioluminiscencia (software Quantity-one®; Bio-Rad). Las concentraciones de los polipéptidos HlyA marcados con E secretados presentes en los sobrenadantes del cultivo de *E. coli* se determinaron por la intensidad de sus bandas de proteínas correspondientes en geles de SDS-poliacrilamida

teñidos con plata [Ansorge, W. 1985. Fast and sensitive detection of protein and DNA bands by treatment with potassium permanganate J. Biochem. Biophys. Methods. 11:13-20] y mediante inmuno-transferencia usando anticuerpo monoclonal anti-E marcado con POD. Se usaron diluciones seriadas de scFvs marcados con E purificados, de concentración conocida, como patrones en estos experimentos.

Entrecruzamiento de proteínas. Antes de su incubación con el agente de entrecruzamiento bifuncional glutarato de disuccinimidilo (DSG, 7.7 Å spacer; Pierce), los polipéptidos de HlyA marcados con E presentes en los sobrenadantes de los cultivos se equilibraron en el mismo volumen de PBS por ultrafiltración a través de una membrana con punto de corte de 10 kDa (Microcon 10, Millipore) que eliminó los compuestos pequeños con grupos amino libres presentes en el medio de cultivo. El entrecruzamiento se llevó a cabo durante 30 minutos a temperatura ambiente añadiendo DBS 40 ó 130 µM a muestras de proteínas de 50 µl equilibradas en PBS. Tras esta incubación, el agente de entrecruzamiento se inactivó con Tris-HCl 50 mM (pH 7,5) durante 15 minutos y se añadió un volumen de tampón de muestra SDS-PAGE 2X concentrado [Fraile, S., F. Roncal, L. A. Fernandez, and V. de Lorenzo 2001. Monitoring Intracellular Levels of XylR in *Pseudomonas putida* with a Single-Chain Antibody Specific for Aromatic-Responsive Enhancer-Binding Proteins J Bacteriol. 183:5571-9]. Tras hervir durante 5 minutos, se cargaron 10 µl para el SDS-PAGE.

Cromatografía por exclusión de tamaño. Los sobrenadantes del cultivo (aproximadamente 0,2 ml) que contenían PBS 1X se mezclaron con 2 mg de proteína patrón de masa conocida (disuelta en 60 µl de H₂O) y se hicieron pasar a través de una resina Bio-Gel A de 1,5m (Bio-Rad) empaquetada en una columna de 1 m de longitud y 1,5 cm de ancho (Bio-Rad). Los patrones de filtración del gel (Bio-Rad) fueron la tiroglobulina (PM 670.000), la gammaglobulina bovina (PM 158.000), la ovoalbúmina de pollo (PM 44.000), la mioglobina equina (PM 17.000) y la vitamina B-12 (PM 1.350). La velocidad de flujo de la muestra a través de la columna se fijó a 0,2 ml/min usando una bomba peristáltica (P-1, Amersham Bioscience). El volumen de vacío de la columna se calculó mediante la elución de azul de dextrano 2000 (Amersham Bioscience). La elución de los patrones de proteínas a través de la columna se monitorizó mediante absorción de luz UV (Uvicord S II, Amersham Bioscience). Se

recogieron fracciones de 1 ml (colector RediFrac, Amersham Bioscience) y se concentraron 10 veces mediante precipitación con ácido tricloroacético (TCA) al 10% (p/v) y 10 µg de albúmina sérica bovina (BSA, Roche) que actuaba como transportador. La presencia de proteínas HlyA marcadas con E en estas fracciones se detectó por
5 Western blot utilizando un anticuerpo monoclonal anti-E marcado con POD (véase más arriba).

Ensayos de inmunoabsorción ligados a enzima (ELISA). Los antígenos (α -amilasa u ovoalbúmina; Sigma) se adsorbieron durante 1 h a 37°C a inmunoplasacas de
10 microtitulación de 96 pocillos (Maxisorb, Nunc) a 200 µg/ml en PBS. Se lavó el exceso de antígeno y las placas se bloquearon durante 16 h a 4°C en tampón MTP (véase anteriormente). Los minianticuerpos se diluyeron en tampón MTP, se añadieron a los pocillos a las concentraciones indicadas en cada caso y se incubaron durante 1 h a temperatura ambiente. A continuación, los anticuerpos no unidos se eliminaron mediante
15 cuatro lavados de los pocillos con PBS que contenía Tween 20 al 0,1% (v/v). El conjugado anticuerpo monoclonal anti-E marcado con POD (0,2 µg/ml en tampón MTP) se añadió a los pocillos y se incubó adicionalmente durante 1 h a temperatura ambiente para detectar los minianticuerpos unidos marcados con E. Tras lavar como antes, las placas se revelaron usando o-fenilenediamina (Sigma). Se dejó continuar la reacción
20 durante 10 minutos en oscuridad, se paró con HCl 0,6 N y se determinó la DO a 490nm de los pocillos (lector de microplacas Benchmark, Bio-Rad).

2. RESULTADOS

Dimerización por ingeniería genética de la señal de secreción de HlyA. En primer
25 lugar, se investigó si la dimerización de C-HlyA podría realizarse por ingeniería genética sin afectar a su secreción. Se seleccionó un corto dominio de unos 6 kDa aproximadamente, denominado ZIP, que se había utilizado para la dimerización de scFvs en el periplasma de *E. coli* [Kerschbaumer, R. J., S. Hirschl, A. Kaufmann, M. Ibl, R. Koenig, and G. Himmler 1997. Single-chain Fv fusion proteins suitable as coating and
30 detecting reagents in a double antibody sandwich enzyme-linked immunosorbent assay Anal Biochem. 249:219-27; Pack, P., and A. Plückthun 1992. Miniantibodies: use of amphipathic helices to produce functional, flexibly linked dimeric FV fragments with high avidity in Escherichia coli Biochemistry. 31:1579-84] para su inserción en C-HlyA.

El dominio ZIP consta de una hélice anfipática que forma la cremallera de leucina del factor de transcripción de levaduras GCN4, flanqueado en su extremo N terminal por una zona bisagra peptídica derivada de la IgG3 del ratón, y en su extremo C terminal por un marcador de polihistidina (6xhis). Se insertó un fragmento de ADN que codificaba para el dominio ZIP internamente cerca del extremo N terminal de la versión de aproximadamente 27 kDa marcada con E de C-HlyA (EHlyA) presente en el plásmido pEHlyA (Figura 1A). El plásmido resultante, pZEHlyA, codifica para un polipéptido de aproximadamente 33 kDa (denominado ZEHlyA) que contiene los dominios ZIP y C-HlyA marcados con E (Figuras 1A y 1B).

La producción de ZEHlyA y EHlyA, como control sin ZIP, se indujo en cultivos de células TolC⁺ de tipo salvaje de *E. coli* (por ejemplo, la cepa HB2151) que llevaban pVDL9.3, que codifica para HlyB y HlyD, y que alojaban pZEHlyA o pEHlyA, respectivamente. Como se muestra en la Figura 1C, se encontraron ambos polipéptidos en niveles similares (aproximadamente 10 µg/ml) en los sobrenadantes de los cultivos de *E. coli* crecidos a 37°C tras la inducción de 4 h con IPTG 0,3 mM. Estas proteínas se detectaron en virtud del péptido marcado con E incorporado en sus secuencias con anticuerpo monoclonal anti-E marcado con POD (Métodos). La secreción de ambos polipéptidos derivados de HlyA fue específica y dependiente de la expresión de los componentes HlyB y HlyD por *E. coli* (datos no mostrados) [Fernández, L. A., and V. De Lorenzo 2001. Formation of disulphide bonds during secretion of proteins through the periplasmic-independent type I pathway Mol Microbiol. 40:332-46]. Este resultado indicaba que la presencia del dominio de dimerización ZIP no tenía efecto en la eficacia de la exportación de la señal de C-HlyA.

A continuación, se investigó el estado de oligomerización de los polipéptidos secretados. Se incubaron muestras de alícuotas que contenía los polipéptidos secretados EHlyA o ZEHlyA con el agente de entrecruzamiento bifuncional glutarato de disuccinimidilo (DSG) y luego se sometieron a SDS-PAGE desnaturante y a inmunotransferencia con anticuerpo monoclonal anti-E marcado con POD (Métodos). En este experimento, sólo se entrecruzaron las muestras de ZEHlyA a baja concentración de DSG (40 µM) para formar una banda de proteínas con una masa molecular (Mr) aparente de aproximadamente 66 kDa (Figura 2, carril 5), lo que estaba de acuerdo con el tamaño esperado para un dímero de ZEHlyA. La concentración más elevada de DSG (130 µM) intensificó la intensidad de la banda correspondiente a ZEHlyA dimerico (Figura 2, carril

6) mientras que sólo tenía una reactividad menor sobre EHlyA control (Figura. 2, carril 3).

También se demostró la dimerización de ZEHlyA mediante cromatografía por exclusión de tamaño. Se separaron muestras de alícuotas de sobrenadantes de cultivos que contenían EHlyA o ZEHlyA secretados en una columna de filtración en gel, con un límite de exclusión de 1.500 kDa, junto con proteínas de masa conocida utilizadas como patrones (Métodos). Como se muestra en la Figura 3A, ZEHlyA mostró una Mr aparente de aproximadamente 66 kDa en la cromatografía de filtración en gel, mientras que EHlyA tuvo una Mr aparente de aproximadamente 32 kDa en las mismas condiciones. Es importante que se detectara un único pico para cada proteína (Figura 3B), lo que indica que ambos polipéptidos estaban presentes como monómeros (EHlyA) y dímeros (ZEHlyA) estables en disolución. Considerados juntos, estos resultados demostraron que podría obtenerse la dimerización de proteínas mediante la incorporación de hélices anfipáticas en C-HlyA sin interferir con su secreción por el transportador de Hly.

Secreción de minianticuerpos diméricos por el sistema de la Hly de *E. coli*. A la vista de los resultados previamente obtenidos se procedió a estudiar si los fragmentos de anticuerpos diméricos podían secretarse mediante el sistema de la Hly. En primer lugar, se construyó un plásmido, denominado pZEHlyA2-SD, para generar fusiones internas entre los fragmentos de anticuerpos recombinantes que carecen del dominio N-SP y ZEHlyA. Este plásmido es un derivado de pEHlyA2-SD [Fernández, L. A., I. Sola, L. Enjuanes, and V. de Lorenzo 2000. Specific secretion of active single-chain Fv antibodies into the supernatants of *Escherichia coli* cultures by use of the hemolysin system Appl Environ Microbiol. 66:5024-5029] en el que se insertó un fragmento de ADN que codifica para el dominio ZIP en un único sitio *SalI*, entre la secuencia de poliunión y el dominio C-HlyA marcado con E (Métodos). Se seleccionaron dos anticuerpos de camello V_{HH}, frente a α -amilasa (*amy*) o la vacuna antitetánica (*ttx*), para determinar su expresión como híbridos con los restos 'EHlyA y 'ZEHlyA (Figuras 4A y 4B). El empleo de anticuerpos de camello V_{HH} como pares de fusión se debió a su pequeño tamaño (aproximadamente 15 kDa) y a su baja tendencia a formar agregados de proteínas [Muyldermans, S. 2001. Single domain camel antibodies: current status J Biotechnol. 74:277-302; Plückthun, A., and P. Pack 1997. New protein engineering approaches to multivalent and bispecific antibody fragments Immunotechnology. 3:83-

105], lo que podría interferir con el análisis de la dimerización obtenida por el dominio ZIP (véase Discusión). Las células de *E. coli* HB2151 (pVDL9.3) se transformaron con un plásmido que codificaba para el híbrido V_{HH} -HlyA (pV_{amy} -HlyA, pV_{amy} -ZHlyA, pV_{itz} -HlyA o pV_{itz} -ZHlyA) y se indujeron 4 h por la adición de IPTG 0,3 mM a los
5 cultivos líquidos crecidos en LB a 30 ó 37 °C. Los polipéptidos secretados V_{amy} -HlyA y V_{amy} -ZHlyA se detectaron posteriormente en los sobrenadantes de los cultivos correspondientes de *E. coli* por Western blot con anticuerpo monoclonal anti-E marcado con POD (Figura 4C). En estas condiciones, la concentración final de V_{amy} -HlyA y V_{amy} -ZHlyA fue de aproximadamente 2 µg/ml a 37°C, y se redujo aproximadamente en 2
10 veces en los cultivos que crecieron a 30°C. Resultados similares se obtuvieron con V_{itz} -HlyA y V_{itz} -ZHlyA (datos no mostrados).

El estado de oligomerización de los híbridos V_{HH} -HlyA secretados se sometió a ensayo mediante cromatografía de filtración en gel (Figura 4D). Muestras de alícuotas de sobrenadantes de cultivos de *E. coli* que contenían V_{amy} -HlyA o V_{amy} -ZHlyA se cargaron
15 en una columna de filtración en gel (límite de exclusión de 1.500 kDa), junto con proteínas de masa conocida. A partir de sus perfiles de elución puede deducirse (Figura 4D) que el híbrido V_{amy} -HlyA tenía una M_r aparente de aproximadamente 40 kDa, lo que estaba completamente de acuerdo con la masa esperada para un monómero de este polipéptido. Por el contrario, V_{amy} -ZHlyA mostró una M_r aparente de aproximadamente
20 95 kDa, que es aproximadamente el doble de la masa esperada de su monómero (es decir, 47 kDa). Hay que mencionar que la temperatura a la que se indujeron los cultivos de *E. coli* (30°C ó 37°C) no tuvo ninguna influencia en el comportamiento cromatográfico de estas muestras. Por lo tanto, la secreción de los anticuerpos de camello V_{HH} monoméricos o diméricos pueden producirse fusionándolos a los restos EHlyA o
25 ZEHlyA, respectivamente.

A continuación, se probó si la dimerización mejoraba las propiedades de unión funcional de V_{amy} -ZHlyA. Para este fin, se comparó la unión a la α -amilasa de V_{amy} -HlyA monomérico y de V_{amy} -ZHlyA dimérico mediante ELISA. En estos experimentos, se incubaron diluciones seriadas de sobrenadantes de cultivos de *E. coli* que contenían
30 cantidades idénticas de V_{amy} -HlyA o V_{amy} -ZHlyA con placas de ELISA recubiertas con α -amilasa u ovoalbúmina (como antígeno control). Tras el lavado, los minianticuerpos unidos se detectaron con el conjugado anticuerpo monoclonal anti-E marcado con POD y se realizó la lectura a la DO_{490nm} (Métodos). La unión específica de la α -amilasa se

demostró incubando estas placas con V_{tx} HlyA y V_{tx} ZHlyA. En la Figura 5 se muestra el resultado de un ELISA prototipo de estos experimentos. La unión previa a ovoalbúmina (en todos los casos la $DO_{490nm} \leq 0,05$) se ha sustraído de los valores presentados. Tal como se indica, la molécula V_{amy} ZHlyA dimérica presentó una afinidad funcional mayor frente a la α -amilasa, que V_{amy} HlyA monomérica. No se observó unión a α -amilasa con los derivados control de V_{tx} (Figura 5) ni con los polipéptidos EHlyA y ZEHlyA (datos no mostrados). En general, al menos se requirió una concentración diez veces mayor de V_{amy} HlyA monomérico para lograr señales de unión a la α -amilasa similares a las obtenidas con V_{amy} ZHlyA. Además, en la concentración de saturación de ambos minianticuerpos (aproximadamente 0,5 μ g/ml) la unión obtenida con V_{amy} ZHlyA alcanzó un nivel meseta superior. Por tanto, la dimerización de V_{amy} ZHlyA induce un efecto de avidéz sobre este minianticuerpo que se refleja en una mayor afinidad de unión funcional a su antígeno.

3. DISCUSIÓN

La dimerización es una propiedad que con frecuencia se desea conseguir por ingeniería genética en las proteínas cuando está implicada una actividad de unión (por ejemplo, en las interacciones proteína-ADN o antígeno-anticuerpo), puesto que puede intensificar su afinidad funcional (avidéz).

Este ejemplo ilustra la obtención por ingeniería genética, por primera vez, de la dimerización de las proteínas secretadas por el sistema de transporte de hemolisina de *E. coli* y se ha empleado esa tecnología para producir minianticuerpos de alta avidéz derivados de los anticuerpos de camello V_{HH} .

Los resultados obtenidos demuestran que la incorporación de una hélice anfipática de autodimerización en el extremo N terminal de C-HlyA no interfiere con la secreción de Hly y permite la dimerización del polipéptido secretado. Tal como se muestra, la dimerización puede intensificar la avidéz de la unión del polipéptido secretado derivado de C-HlyA. Además, también puede tener otras aplicaciones, como la asociación molecular de varios antígenos y/o adyuvantes producidos por cepas bacterianas vivas, o la combinación de diversas actividades biológicas para la generación de moléculas biespecíficas (por ejemplo, la unión a un antígeno y el reclutamiento del complemento).

scFvs diméricos de alta avidez se han producido en el periplasma de las células de *E. coli* insertando hélices anfipáticas en su extremo C terminal. Debido a la tendencia de algunos scFvs a formar dímeros y agregados de proteínas de elevado peso molecular se utilizaron fragmentos más pequeños de anticuerpos.

- 5 Los anticuerpos de camello V_{HH} han recibido mucha atención debido a su mejor solubilidad y a su estructura más simple, lo que facilita su amplificación y clonaje. Merece la pena destacar que los cambios efectuados no disminuyen la afinidad ni la especificidad de los anticuerpos de camello V_{HH} , debido a la presencia de regiones determinantes de la complementaridad (CDR) extremadamente variables que compensan
- 10 la pérdida de diversidad provocada por la ausencia de un dominio V_L . Los anticuerpos de camello también han demostrado un potencial extraordinario como inhibidores enzimáticos puesto que sus grandes CDR pueden alcanzar los sitios activos escondidos en las enzimas. Además, la similitud entre los anticuerpos de camello V_{HH} y las secuencias de la familia VH3 humana está permitiendo la generación de librerías
- 15 presentadas en fagos de los dominios V_H humanos camelizados y de los anticuerpos de camello V_{HH} humanizados.

- Los beneficios expuestos anteriormente motivaron a los inventores a utilizar los dominios V_{HH} para su secreción por el translocador de hemolisina de *E. coli*. Los resultados obtenidos muestran que los anticuerpos de camello funcionales, tanto en la
- 20 forma monomérica como en la dimérica, pueden recuperarse de los sobrenadantes del cultivo de *E. coli* a niveles similares a los obtenidos en su expresión periplásmica (aproximadamente 1 mg/litro de cultivo a D.O._{600nm} =1). Además, la dimerización provocada por ZEHlyA indujo un aumento de diez veces en la afinidad funcional de V_{amy} . Este valor está dentro del intervalo esperado producido por el cambio de
- 25 anticuerpos monovalentes a divalentes. En conclusión, estos datos demuestran que el sistema de secreción de Hly puede emplearse para la secreción de polipéptidos y minianticuerpos diméricos de alta avidez. La simplicidad de esta tecnología puede ser extremadamente útil para la selección de alto rendimiento de las librerías de anticuerpos.

REIVINDICACIONES

1. Una construcción de ADN que comprende:

- a) una primera secuencia de ácido nucleico que contiene la secuencia de nucleótidos que codifica para un producto de interés;
- 5 b) una segunda secuencia de ácido nucleico que contiene la secuencia de nucleótidos que codifica para un dominio de dimerización; y
- c) una tercera secuencia de ácido nucleico que contiene la secuencia de nucleótidos que codifica para la α -hemolisina (HlyA) de *E. coli* o para un fragmento de dicha proteína que comprende la señal de reconocimiento del mecanismo de secreción del sistema transportador de Hly de *E. coli*, o una
10 secuencia de nucleótidos que codifica para un gen homólogo, o una secuencia de nucleótidos que codifica para una variante, natural o artificial, de HlyA o de un fragmento de la misma que comprende la señal de reconocimiento del mecanismo de secreción del sistema transportador de Hly de *E. coli*.

15 2.- Una construcción de ADN según la reivindicación 1, en donde el extremo 3' de dicha primera secuencia de ácido nucleico está unido al extremo 5' de dicha segunda secuencia de ácido nucleico y el extremo 3' de dicha segunda secuencia de ácido nucleico está unido al extremo 5' de dicha tercera secuencia de ácido nucleico.

20 3. Construcción de ADN según la reivindicación 1, en la que dicho producto de interés se selecciona entre enzimas, inhibidores enzimáticos, hormonas, moléculas implicadas en la adhesión y/o señalización celular, moléculas implicadas en detección o marcaje, moléculas compuestas por dominios, antígenos inmunogénicos, agentes
25 terapéuticos, y moléculas inmunoreguladoras.

30 4. Construcción de ADN según la reivindicación 3, en la que dicho producto de interés se selecciona entre antígenos específicos de tumor, antígenos de enfermedades auto-inmune, factores de crecimiento, citoquinas, interleuquinas, interferones y minianticuerpos.

5. Construcción de ADN según la reivindicación 1, en la que dicho dominio de dimerización comprende una hélice peptídica o una estructura de doble arrollamiento helicoidal (coiled coil).

5 6. Construcción de ADN según la reivindicación 5, en la que dicho dominio de dimerización comprende una cremallera de leucina.

7. Construcción de ADN según la reivindicación 5, en la que dicho dominio de dimerización comprende la cremallera de leucina del factor de transcripción GCN4 de
10 levadura.

8. Construcción de ADN según la reivindicación 1, en la que dicha tercera secuencia de ácido nucleico se selecciona entre:

- a) la secuencia de nucleótidos que codifica para la HlyA de *E. coli*;
- 15 b) una secuencia de ácido nucleico que comprende la secuencia de nucleótidos que codifica para los 60 últimos aminoácidos del extremo C-terminal de HlyA de *E. coli*;
- c) una secuencia de ácido nucleico constituida por una secuencia de nucleótidos que codifica para los 60 últimos aminoácidos del extremo C-terminal de HlyA de *E. coli*;
- 20 d) la secuencia de nucleótidos identificada como SEQ ID NO: 1; y
- e) una secuencia de nucleótidos que codifica para la secuencia de aminoácidos identificada como SEQ ID NO: 2.

25 9. Construcción de ADN según las reivindicaciones 1 y 2, que comprende, además, una cuarta secuencia de ácido nucleico que codifica para un péptido espaciador situada entre dichas primera y segunda secuencias de ácido nucleico, en donde el extremo 5' de dicha cuarta secuencia de ácido nucleico está unido al extremo 3' de dicha primera secuencia de ácido nucleico y el extremo 3' de dicha cuarta secuencia de ácido nucleico
30 está unido al extremo 5' de dicha segunda secuencia de ácido nucleico.

10. Construcción de ADN según la reivindicación 9, en la que dicho péptido espaciador comprende repeticiones de restos de aminoácidos, preferentemente, repeticiones Gly-Gly-Gly-Ser.

5 11. Construcción de ADN según la reivindicación 9, en el que dicho péptido espaciador es una región bisagra de un anticuerpo.

10 12. Construcción de ADN según la reivindicación 1, que comprende, además, una quinta secuencia de ácido nucleico que codifica para un péptido susceptible de ser utilizado con fines de aislamiento o purificación.

15 13. Construcción de ADN según la reivindicación 12, en la que dicho péptido susceptible de ser utilizado con fines de aislamiento o purificación comprende una secuencia de polihistidina o una secuencia peptídica reconocida por un anticuerpo monoclonal y que puede servir para purificar la proteína de fusión resultante por cromatografía de inmunoafinidad.

20 14. Construcción de ADN según la reivindicación 12, en la que dicha quinta secuencia de ácido nucleico está situada entre dichas segunda y tercera secuencias de ácido nucleico ordenadas según la reivindicación 2, en donde el extremo 5' de dicha quinta secuencia de ácido nucleico está unido al extremo 3' de dicha segunda secuencia de ácido nucleico y el extremo 3' de dicha quinta secuencia de ácido nucleico está unido al extremo 5' de dicha tercera secuencia de ácido nucleico.

25 15. Construcción de ADN según la reivindicación 1, que comprende, además, una sexta secuencia de ácido nucleico que codifica para un péptido susceptible de ser utilizado con fines de reconocimiento.

30 16. Construcción de ADN según la reivindicación 15, en la que dicho péptido susceptible de ser utilizado con fines de reconocimiento comprende una secuencia peptídica reconocida por un anticuerpo monoclonal y que puede servir para reconocer la proteína de fusión resultante por técnicas de inmunodetección.

17. Construcción de ADN según la reivindicación 15, en la que dicho péptido susceptible de ser utilizado con fines de reconocimiento comprende la secuencia del epítipo E.

5 18. Construcción de ADN según la reivindicación 15, en la que dicha sexta secuencia de ácido nucleico está situada entre dichas segunda y tercera secuencias de ácido nucleico ordenadas según la reivindicación 2, en donde el extremo 5' de dicha sexta secuencia de ácido nucleico está unido al extremo 3' de dicha segunda secuencia de ácido nucleico y el extremo 3' de dicha sexta secuencia de ácido nucleico está unido al extremo
10 5' de dicha tercera secuencia de ácido nucleico.

19. Construcción de ADN según las reivindicaciones 12 ó 15, en la que dichas quinta y sexta secuencias de ácido nucleico están separadas entre sí.

15 20. Construcción de ADN según las reivindicaciones 12 ó 15, en la que dichas quinta y sexta secuencias de ácido nucleico están unidas entre sí.

21.- Construcción de ADN según la reivindicación 20, en la que dichas quinta y sexta secuencias de ácido nucleico están situadas entre las secuencias segunda y tercera
20 ordenadas según la reivindicación 2.

22. Construcción de ADN según la reivindicación 1, que comprende, además, una séptima secuencia de ácido nucleico que comprende una secuencia de nucleótidos que codifica para una secuencia de aminoácidos susceptible de ser escindida específicamente
25 por medios enzimáticos o químicos.

23. Construcción de ADN según la reivindicación 22, en la que dicha séptima secuencia de ácido nucleico comprende una secuencia de nucleótidos que codifica para un sitio de reconocimiento de una proteasa.
30

24. Construcción de ADN según la reivindicación 23, en la que dicha proteasa se selecciona entre una enteroquinasa, Arg-C endoproteasa, Glu-C endoproteasa, Lys-C endoproteasa y Factor de coagulación Xa.

25. Construcción de ADN según la reivindicación 22, en la que dicha séptima secuencia de ácido nucleico comprende una secuencia de nucleótidos que codifica para un sitio susceptible de ser específicamente escindido por un reactivo químico.

5 26. Construcción de ADN según la reivindicación 25, en la que dicho reactivo químico es bromuro de cianógeno.

27. Construcción de ADN según la reivindicación 22, en la que dicha séptima secuencia de ácido nucleico se encuentra en cualquier posición entre las secuencias
10 segunda y tercera de ácido nucleico ordenadas según la reivindicación 2.

28. Un cassette de expresión que comprende una construcción de ADN según cualquiera de las reivindicaciones anteriores operativamente unida a una secuencia de control de expresión.
15

29. Cassette de expresión según la reivindicación 28, en la que dicha secuencia de control de expresión comprende una secuencia promotora, una secuencia codificante para reguladores transcripcionales, una secuencia de unión a ribosomas (RBS) y/o una secuencia terminadora de transcripción.
20

30. Cassette de expresión según las reivindicaciones 28 y 29, en el que dicha secuencia de control de expresión es funcional en bacterias.

31. Vector que comprende al menos una construcción de ADN según cualquiera
25 de las reivindicaciones 1 a 27 o al menos un cassette de expresión según cualquiera de las reivindicaciones 28 a 30.

32. Vector según la reivindicación 31, que comprende, además, un marcador.

30 33. Vector según la reivindicación 32, en el que dicho marcador comprende un gen de resistencia a antibióticos o un gen de resistencia a compuestos tóxicos.

34. Una bacteria Gram negativa que comprende una construcción de ADN según

cualquiera de las reivindicaciones 1 a 27, o un cassette de expresión según cualquiera de las reivindicaciones 28 a 30.

35. Una bacteria Gram negativa que comprende al menos una construcción de ADN según cualquiera de las reivindicaciones 1 a 27 o al menos un cassette de expresión según cualquiera de las reivindicaciones 28 a 30 o al menos un vector según las reivindicaciones 31 a 33.

36. Bacteria según la reivindicación 35, seleccionada entre *Escherichia coli*, *Salmonella tiphyrium*, *Pseudomonas aeruginosa* y *Pseudomonas putida*.

37. Una proteína de fusión dimérica obtenible por expresión de la secuencias de ácido nucleico contenida en una construcción de ADN según cualquiera de las reivindicaciones 1 a 27.

38. Una proteína de fusión dimérica obtenible por expresión de la secuencias de ácido nucleico contenidas en al menos una construcción de ADN según cualquiera de las reivindicaciones 1 a 27, o al menos un cassette de expresión según cualquiera de las reivindicaciones 28 a 30 o al menos un vector según las reivindicaciones 31 a 33.

39. Proteína de fusión dimérica según las reivindicaciones 37 y 38, en la que cada monómero comprende:

- (i) la secuencia de aminoácidos de un producto de interés.
- (ii) una secuencia de aminoácidos correspondiente a un dominio de dimerización; y
- (iii) la secuencia de aminoácidos de la α -hemolisina (HlyA) de *Escherichia coli* o de un fragmento de dicha proteína que comprende la señal de reconocimiento del mecanismo de secreción del sistema transportador de hemolisina (Hly) de *E. coli*;

40. Proteína de fusión según la reivindicación 40, en la que cada monómero comprende:

- 5 (i) un producto de interés seleccionado entre una enzima, un inhibidor enzimático, una hormona, una molécula implicada en la adhesión y/o señalización celular, moléculas implicadas en detección o marcaje, moléculas compuestas por dominios, un antígeno inmunogénico, un agente terapéutico, una molécula inmunoreguladora;
- (ii) un dominio de dimerización, seleccionado entre una hélice peptídica y una estructura de doble arrollamiento helicoidal (coiled coil); y
- 10 (iii) la secuencia completa de aminoácidos de HlyA de *E. coli*, o un fragmento de HlyA de *E. coli* que comprende la señal de reconocimiento del mecanismo de secreción del sistema transportador de Hly de *E. coli*.

15 41. Proteína de fusión según la reivindicación 41, en la que cada monómero comprende:

- (i) un producto de interés seleccionado entre un antígeno específico de tumor, un antígeno de enfermedades autoinmunes, un factor de crecimiento, una citoquina, una interleuquina, un interferón, y un minianticuerpo;
- 20 (ii) un dominio de dimerización, seleccionado entre una hélice peptídica y una estructura de doble arrollamiento helicoidal (coiled coil); y
- (iii) la secuencia completa de aminoácidos de HlyA de *E. coli*, o un fragmento de HlyA de *E. coli* que comprende la señal de reconocimiento del mecanismo de secreción del sistema transportador de Hly de *E. coli*.

25

42. Proteína de fusión según las reivindicaciones 37 a 41, en la que cada monómero comprende además (a) un péptido espaciador entre el producto de interés y el dominio de dimerización; y/o (b) un péptido para facilitar el aislamiento o purificación del péptido o proteína de fusión; y/o (c) un péptido que permita el reconocimiento del péptido o proteína de fusión; y/o (d) una secuencia de aminoácidos susceptible de ser escindida específicamente por medios enzimáticos o químicos.

30

43. Proteína de fusión según la reivindicación 42, en la que cada monómero

comprende, (a) un péptido que contiene repeticiones de Gly-Gly-Gly-Ser o la región bisagra de un anticuerpo; y/o (b) una secuencia de polihistidina o una secuencia peptídica reconocida por un anticuerpo monoclonal y que puede servir para purificar la proteína de fusión resultante por cromatografía de inmunoafinidad; y/o (c) una
5 secuencia peptídica reconocida por un anticuerpo monoclonal y que puede servir para reconocer la proteína de fusión resultante por técnicas de inmunodetección; y/o (d) una secuencia de aminoácidos que constituye un sitio de reconocimiento de una enteroquinasa, Arg-C endoproteasa, Glu-C endoproteasa, Lys-C endoproteasa o Factor de coagulación Xa o una secuencia de aminoácidos susceptible de ser específicamente
10 escindida por un reactivo químico.

44. Un método para producir un producto de interés, en forma de una proteína de fusión dimérica según cualquiera de las reivindicaciones 37 a 43, que comprende crecer una bacteria según la reivindicaciones 34 a 36 bajo condiciones que permiten la
15 producción y excreción al medio de cultivo de dicho producto de interés en forma de una proteína de fusión dimérica.

45. Un método según la reivindicación 44 para producir una proteína de fusión dimérica que comprende dos productos de interés.

20

46. Método según la reivindicaciones 44 y 45, que comprende, además, el aislamiento y purificación de dicha proteína de fusión dimérica.

47. Uso de una construcción de ADN según las reivindicaciones 1 a 27 para la
25 creación de una librería de proteínas diméricas.

48. Uso de la librería de la reivindicación anterior para la selección de moléculas con capacidad de unión a un antígeno determinado.

30 49. Proteína de fusión dimérica según las reivindicaciones 37 a 43 para uso en terapia.

50. Proteína de fusión dimérica según las reivindicaciones 37 a 43 para uso en diagnóstico *in vitro* o *in vivo*.

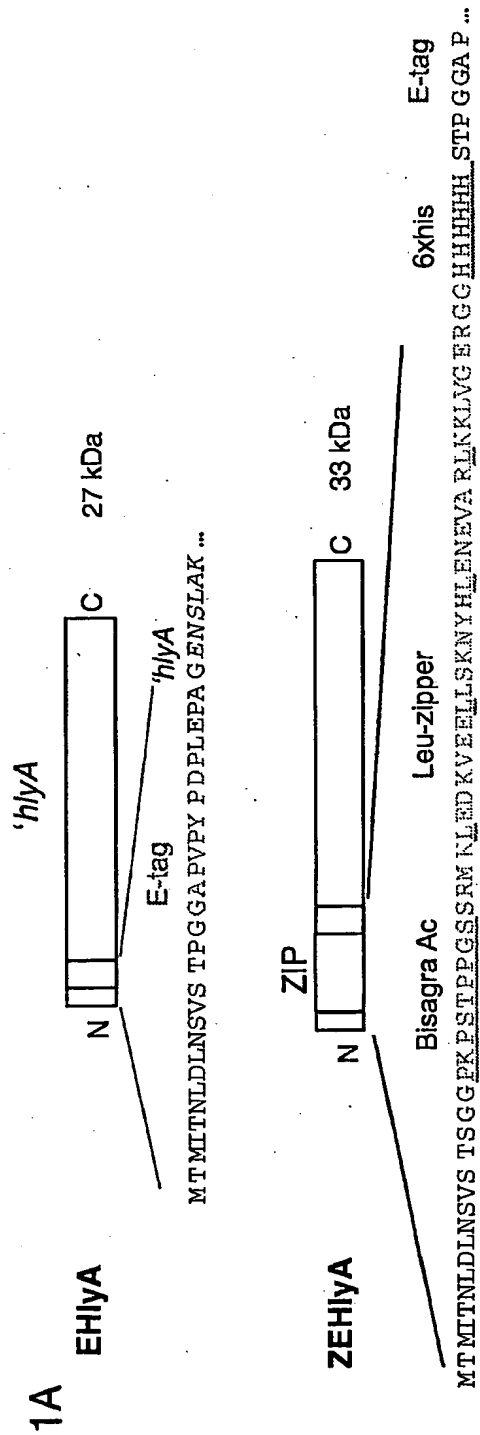


Fig. 1 (cont.)

2/30

1B

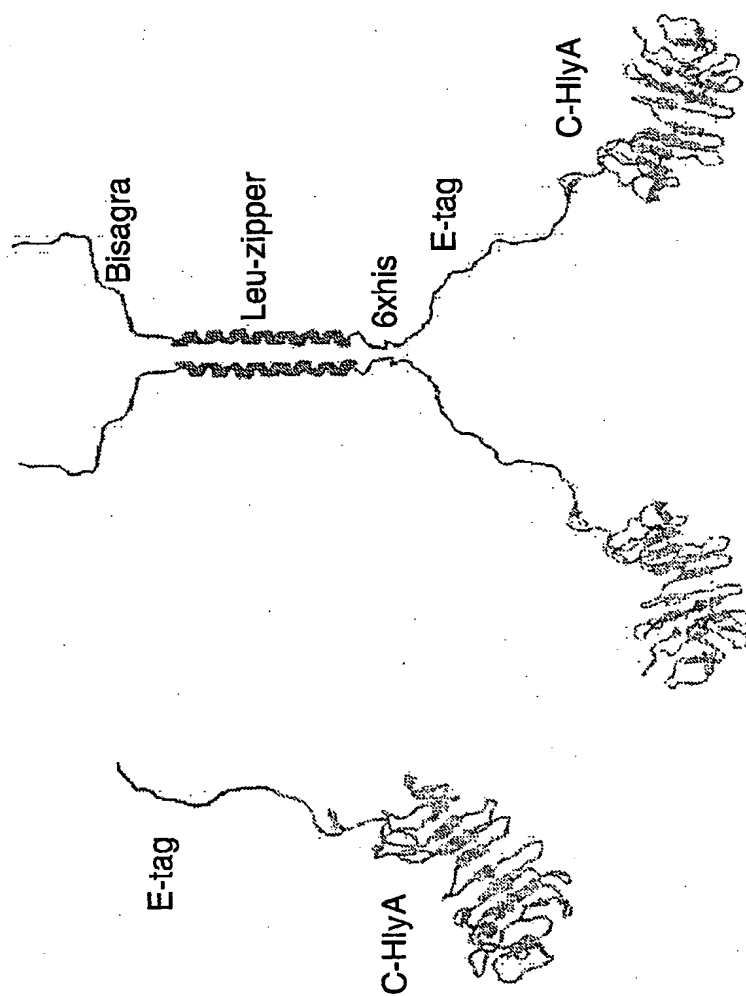


Fig. 1 (cont.)

3/30

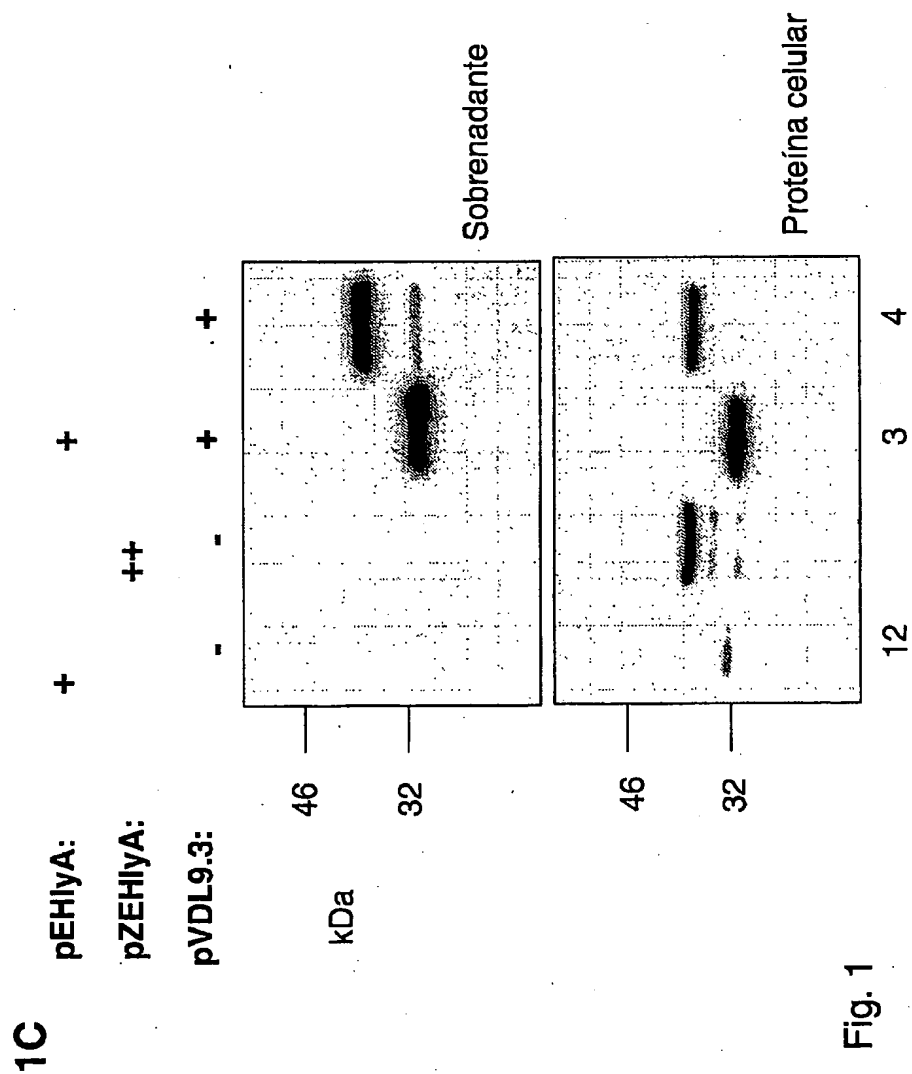


Fig. 1

4/30

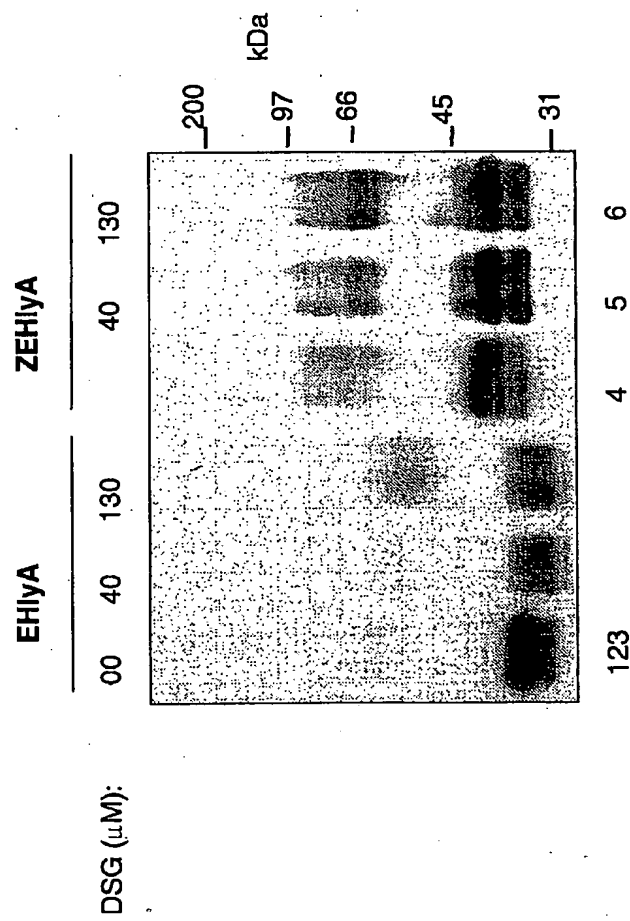


Fig. 2

5/30

3A

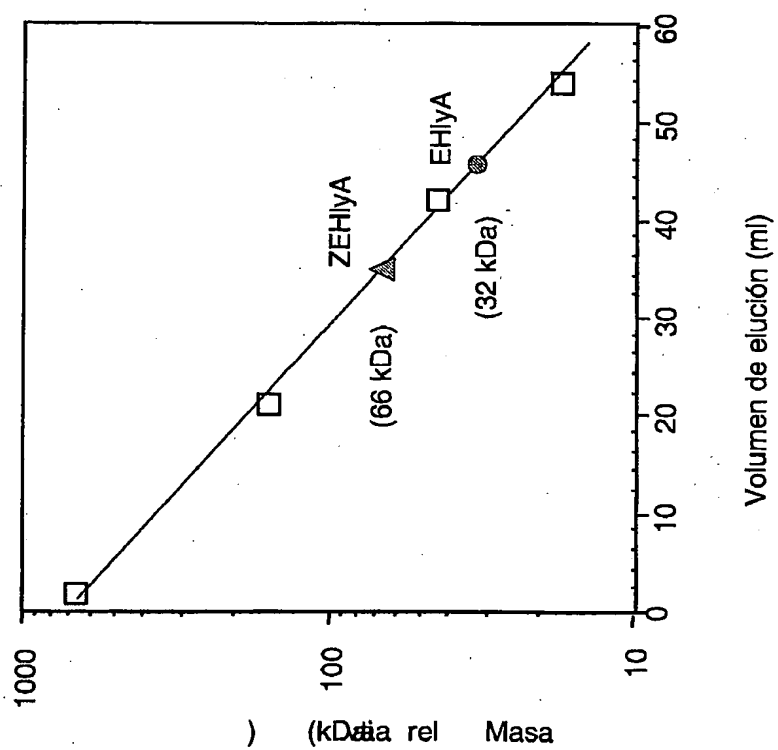
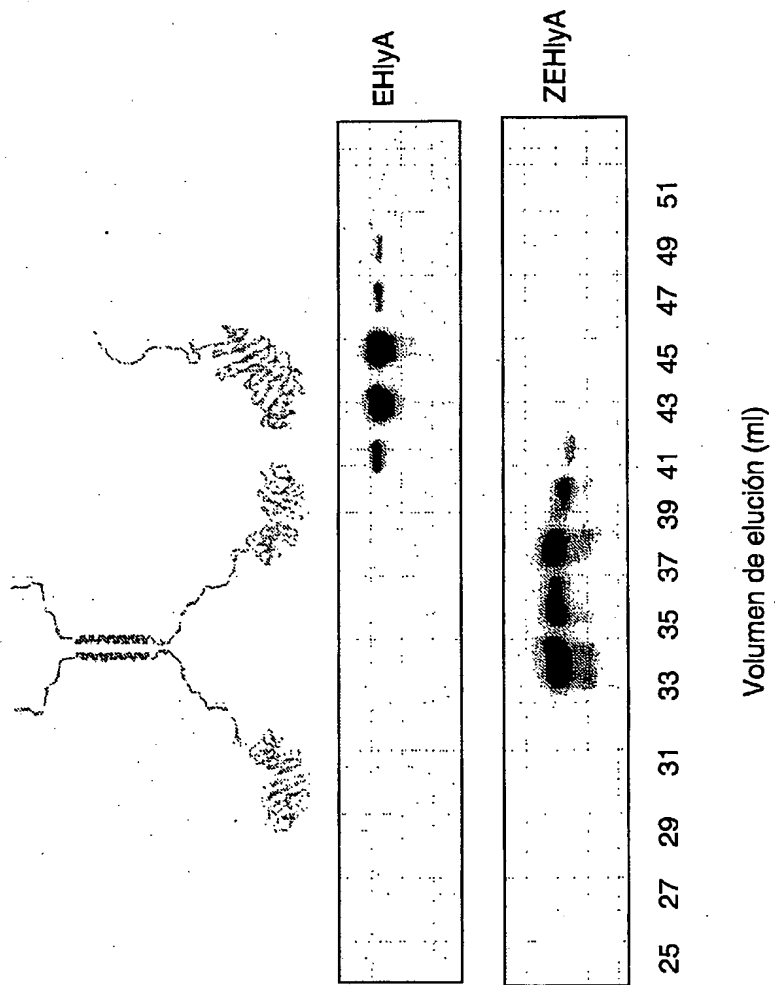


Fig. 3 (cont.)

6/30

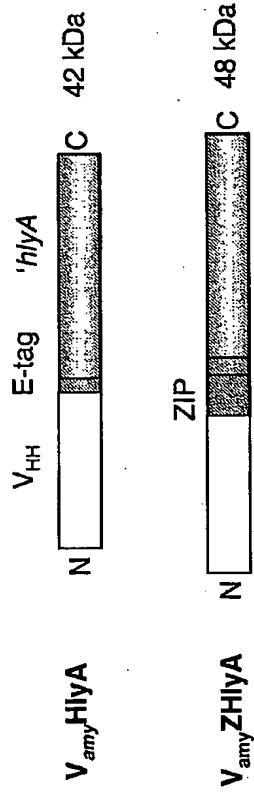
3B

Fig. 3



7/30

4A



4B

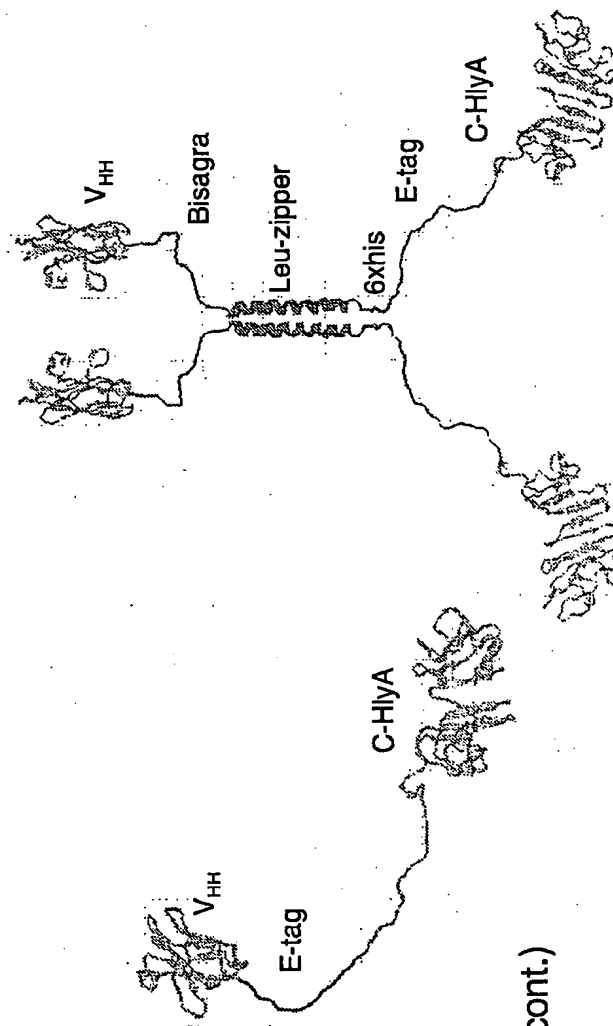


Fig. 4 (cont.)

4C

4D

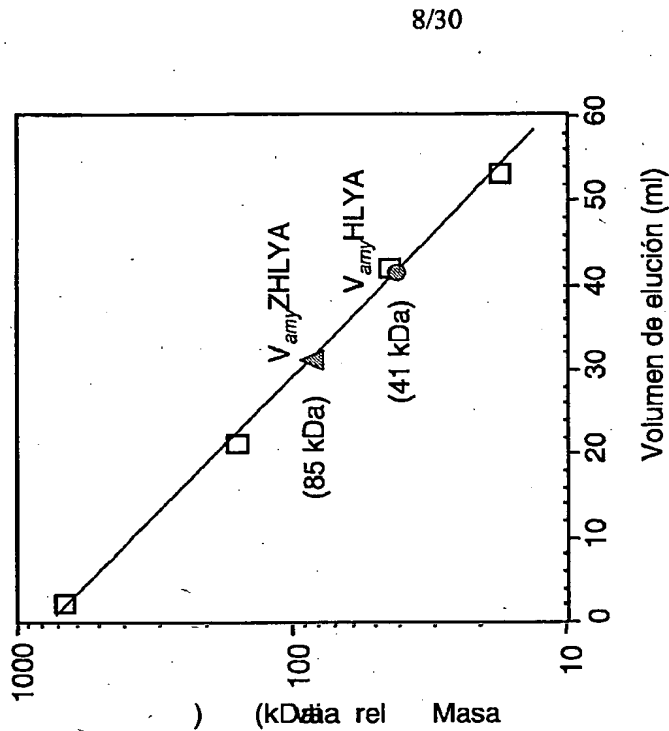
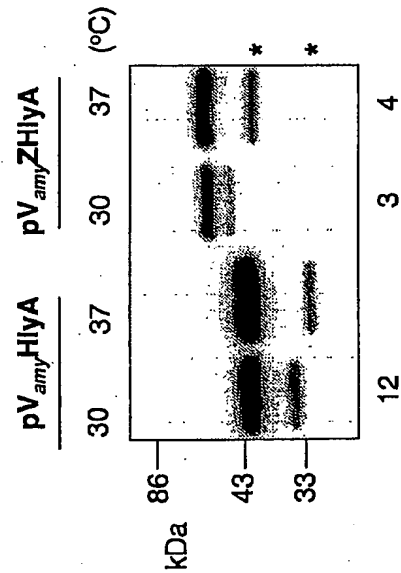
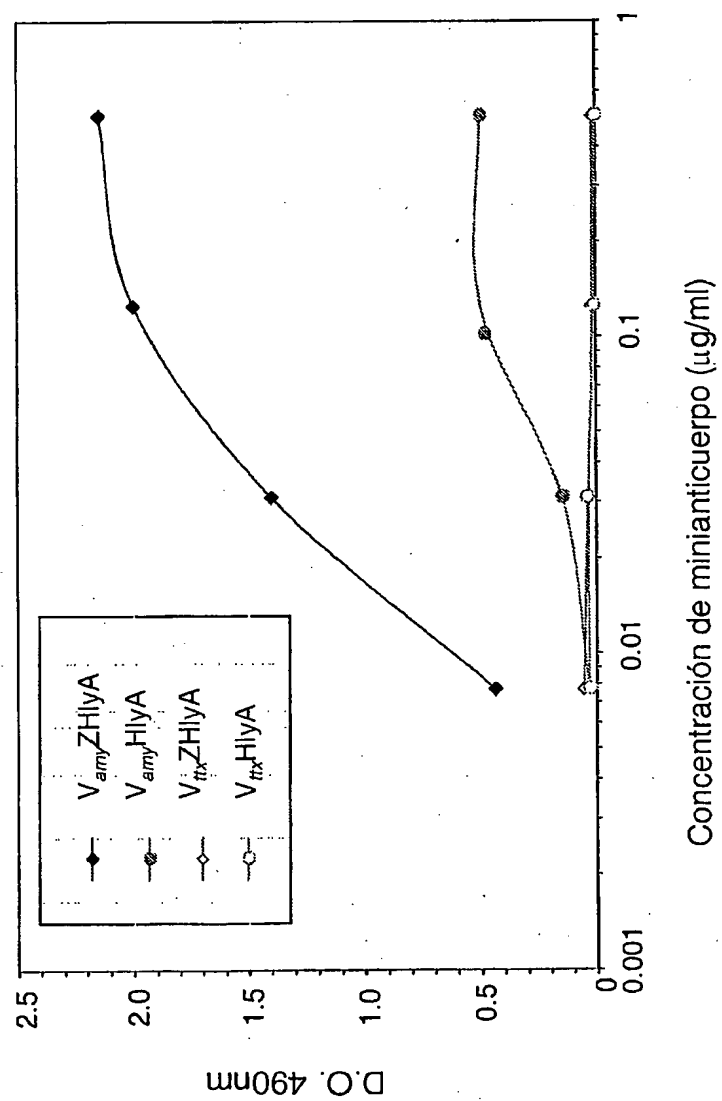


Fig. 4

9/30

Fig. 5



Mapa de pZEHlyA

AGCGGATAACAATTTACACAGGAAACAGCT (secuencia anterior a ATG)

Fig. 6 (cont.)

11/30

GGTAATGACAAGTTGTACGGCAGTGAGGGAGCAGACCTGCTTGATGGCGGAGAAGGGAAT
301 -----+-----+-----+-----+-----+ 360
CCATTACTGTTCAACATGCCGTCACCTCCCTCGTCTGGACGAACTACCGCCTCTTCCCTTA
a G N D K L Y G S E G A D L L D G G E G N -

GATCTTCTGAAAGGTGGATATGGTAATGATATTTATCGTTATCTTTCAGGATATGGCCAT
361 -----+-----+-----+-----+-----+ 420
CTAGAAGACTTTCCACCTATACCATTACTATAAATAGCAATAGAAAGTCCTATACCGGTA
a D L L K G G Y G N D I Y R Y L S G Y G H -

CATATTATTGACGATGAAGGGGGGAAAGACGATAAACTCAGTTTAGCTGATATAGATTTT
421 -----+-----+-----+-----+-----+ 480
GTATAATAACTGCTACTTCCCCCTTTCTGCTATTTGAGTCAAATCGACTATATCTAAAG
a H I I D D E G G K D D K L S L A D I D F -

CGGGACGTTGCCTTTAAGCGAGAAGGGAATGACCTCATTATGTATAAAGCTGAAGGTAAT
481 -----+-----+-----+-----+-----+ 540
GCCCTGCAACGGAAATTCGCTCTTCCCTTACTGGAGTAATACATATTTGCACTTCCATTA
a R D V A F K R E G N D L I M Y K A E G N -

GTTCTTTCTATTGGCCACAAAAATGGTATTACATTTAAAACTGGTTTGAAAAAGAGTCA
541 -----+-----+-----+-----+-----+ 600
CAAGAAAGATAACCGGTGTTTTTACCATAATGTAAATTTTGACCAAACTTTTTCTCAGT
a V L S I G H K N G I T F K N W F E K E S -

GATGATCTCTCTAATCATCAGATAGAGCAGATTTTGTATAAAGACGGCAGGGTAATCACA
601 -----+-----+-----+-----+-----+ 660
CTACTAGAGAGATTAGTAGTCTATCTCGTCTAAAACTATTTCTGCCGTCCCATTAGTGT
a D D L S N H Q I E Q I F D K D G R V I T -

CCAGATTCTCTTAAAAAGCATTGGAATATCAGCAGAGTAATAACAAGGTAAGTTATGTG
661 -----+-----+-----+-----+-----+ 720
GGTCTAAGAGAATTTTTTCGTAACTTATAGTCGTCTCATTATGTTCCATTCAATACAC
a P D S L K K A F E Y Q Q S N N K V S Y V -

TATGGACATGATGCATCAACTTATGGGAGCCAGGACAATCTTAATCCATTAATTAATGAA
721 -----+-----+-----+-----+-----+ 780
ATACCTGTACTACGTAGTTGAATACCCTCGGTCTGTAGAAATTAGGTAATTAATTACTT
a Y G H D A S T Y G S Q D N L N P L I N E -

Fig. 6 (cont.)

12/30

PstI
|

781 ATCAGCAAAATCATTTCAGCTGCAGGTAACCTTCGATGTTAAGGAGGAAAGATCTGCCGCT 840
 -----+-----+-----+-----+-----+
 TAGTCGTTTTAGTAAAGTCGACGTCCATTGAAGCTACAATTCCTCCTTCTAGACGGCGA

a I S K I I S A A G N F D V K E E R S A A -

NdeI
|

841 TCTTTATTGCAGTTGTCCGGTAATGCCAGTGATTTTTCATATGGACGGAACCTCAATAACT 900
 -----+-----+-----+-----+-----+
 AGAAATAACGTCAACAGGCCATTACGGTCACTAAAAAGTATACCTGCCTTGAGTTATTGA

a S L L Q L S G N A S D F S Y G R N S I T -

TTGACAGCATCAGCATAA

901 -----+----- 918
 AACTGTCGTAGTCGTATT

a L T A S A * -

Enzimas que cortan:

AccI	BamHI	EcoRI	HindIII	NdeI	PstI	SacI	Sall
SmaI	XbaI						

Enzimsa que no cortan:

KpnI	NcoI	NheI	NotI	SacII	SphI
------	------	------	------	-------	------

Fig. 6

13/30

Mapa de pZEHLA2SD

Con 11 enzimas: ECORI BGLII BAMHI NCOI NHEI SFII SALI XMAI XBAI EAGI SALI
HINDIII

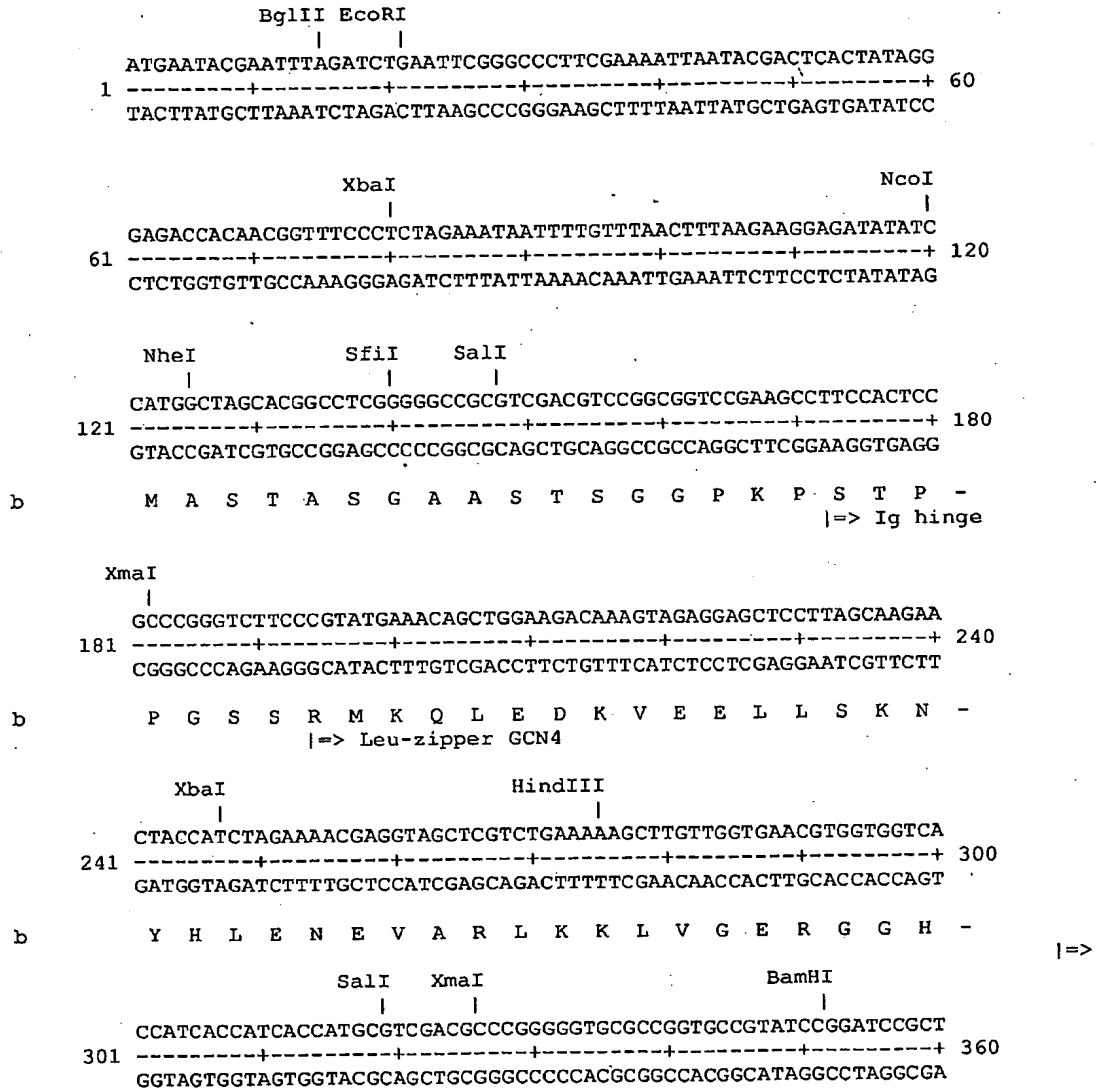


Fig. 7 (cont.)

14/30

b H H H H H A S T P G G A P V P Y P D P L -
 6xhis tag | => E-tag
 EagI
 |
GGAACCGCGCCGGGAAAATTCTCTTGCTAAAAATGTATTATCCGGTGGAAAAGGTAATGA
361 -----+-----+-----+-----+-----+ 420
CCTTGGCCGGCCCTTTTAAGAGAACGATTTTACATAATAGCCACCTTTTCCATTACT

b E P A G E N S L A K N V L S G G K G N D -
 | => C-hlyA
CAAGTTGTACGGCAGTGAGGGAGCAGACCTGCTTGATGGCGGAGAAGGGAATGATCTTCT
421 -----+-----+-----+-----+-----+ 480
GTTCAACATGCCGTCACCTCCCTCGTCTGGACGAACCTACCGCCTCTTCCCTTACTAGAAGA

b K L Y G S E G A D L L D G G E G N D L L -
GAAAGGTGGATATGGTAATGATATTTATCGTTATCTTTTCCAGGATATGGCCATCATATTAT
481 -----+-----+-----+-----+-----+ 540
CTTCCACCTATACCATTACTATAAATAGCAATAGAAAGTCTATACCGGTAGTATAATA

b K G G Y G N D I Y R Y L S G Y G H H I I -
TGACGATGAAGGGGGGAAAGACGATAAACTCAGTTTAGCTGATATAGATTTCCGGGACGT
541 -----+-----+-----+-----+-----+ 600
ACTGCTACTTCCCCCTTTCTGCTATTTGAGTCAAATCGACTATATCTAAAGGCCCTGCA

b D D E G G K D D K L S L A D I D F R D V -
TGCCCTTTAAGCGAGAAGGGAATGACCTCATTATGTATAAAGCTGAAGGTAATGTTCTTTC
601 -----+-----+-----+-----+-----+ 660
ACGGAAATTCGCTCTTCCCTTACTGGAGTAATACATATTTGACTTCCATTACAAGAAAG

b A F K R E G N D L I M Y K A E G N V L S -
TATTGGCCACAAAATGGTATTACATTTAAAAACTGGTTTGAAAAAGAGTCAGATGATCT
661 -----+-----+-----+-----+-----+ 720
ATAACCGGTGTTTTTACCATAATGTAATTTTGACCAAACCTTTTCTCAGTCTACTAGA

b I G H K N G I T F K N W F E K E S D D L -
CTCTAATCATCAGATAGAGCAGATTTTGTATAAAGACGGCAGGTAATCACACCAGATTC
721 -----+-----+-----+-----+-----+ 780
GAGATTAGTAGTCTATCTCGTCTAAAACTATTTCTGCCGTCCCATAGTGTGGTCTAAG

b S N H Q I E Q I F D K D G R V I T P D S -
TCTTAAAAAGCATTTGAATATCAGCAGAGTAATAACAAGGTAAGTTATGTGTATGGACA
781 -----+-----+-----+-----+-----+ 840
AGAATTTTTTCGTAACTTATAGTCGTCTCATTATTGTTCCATTCAATACACATACCTGT

Fig. 7 (cont.)

15/30

b L K K A F E Y Q Q S N N K V S Y V Y G H -
 TGATGCATCAACTTATGGGAGCCAGGACAATCTTAATCCATTAATTAATGAAATCAGCAA
 841 -----+-----+-----+-----+-----+ 900
 ACTACGTAGTTGAATACCTCGGTCTGTTAGAATTAGGTAATTAATTACTTTAGTCGTT

b D A S T Y G S Q D N L N P L I N E I S K -
 BglII
 |
 AATCATTTCAGCTGCAGGTAACCTTCGATGTTAAGGAGGAAAGATCTGCCGCTTCTTTATT
 901 -----+-----+-----+-----+-----+ 960
 TTAGTAAAGTCGACGTCCATTGAAGCTACAATTCCTCCTTTCTAGACGGCGAAGAAATAA

b I I S A A G N F D V K E E R S A A S L L -
 GCAGTTGTCCGGTAATGCCAGTGATTTTTCATATGGACGGAACCTCAATACTTTGACAGC
 961 -----+-----+-----+-----+-----+ 1020
 CGTCAACAGGCCATTACGGTCACTAAAAAGTATACCTGCCTTGAGTTATTGAAACTGTCTG

b Q L S G N A S D F S Y G R N S I T L T A -
 ATCAGCATAATATATTAATTTAAATGATAGCAATCTTACTGGGCTGTGCCACATAAGATT
 1021 -----+-----+-----+-----+-----+ 1080
 TAGTCGTATTATATAATTAATTTACTATCGTTAGAATGACCCGACACGGTGTATTCTAA

b S A * -
 GCTATTTTTTTGGAGTCATAATGGATTCTTGTCATAAAATTGATTATGGGTTATACGCC
 1081 -----+-----+-----+-----+-----+ 1140
 CGATAAAAAACCTCAGTATTACCTAAGAACAGTATTTTAACTAATACCCAATATGCGGG

1141 TGGAGATTTTAGCCCAATACCATAACGTCTCTGTTAACCCGGAAGAAATTAACATAGAT
 -----+-----+-----+-----+-----+ 1200
 ACCTCTAAAATCGGGTTATGGTATTGCAGAGACAATTGGGCCTTCTTTAATTTGTATCTA

1201 TTGACACAGACGGGACTGGTCTGGGATTAACGTCTGCTTGCTTGCTGCGAAATCTTTAG
 -----+-----+-----+-----+-----+ 1260
 AACTGTGTCTGCCCTGACCAGACCTAATTGCAGTACCAACGAACGACGCTTTAGAAATC

1261 AACTAAAGGTAAAACAGGTAAAAAAACAATTGACCGATTAACTTTATTTCTCTGCCCG
 -----+-----+-----+-----+-----+ 1320
 TTGATTTCCATTTTGTCCATTTTTTTTGTAACTGGCTAATTGAAATAAGAGACGGGC

1321 CATTAGTCTGGAGAGAGGATGGACGTCATTTTATTCTGACTAAAGTCAGTAAAGAAGCAA
 -----+-----+-----+-----+-----+ 1380
 GTAATCAGACCTCTCTCTACCTGCAGTAAAAATAAGACTGATTCAGTCATTTCTTCGTT

Fig. 7 (cont.)

16/30

```

1381 ACAGATATCTTATTCTGATCTGGAGCAGCGAAATCCCGTGTTCTCGAACAGTCTGAGT
-----+-----+-----+-----+-----+-----+ 1440
TGTCTATAGAATAAAGACTAGACCTCGTCGCTTAGGGGCACAAGAGCTTGTCTCAGACTCA

1441 TTGAGGCGTTATATCAGGGGCATATTATTCTTATCGCTTCCCGTTCTTCTGTTGCCGGGA
-----+-----+-----+-----+-----+-----+ 1500
AACTCCGCAATATAGTCCCGGTATAATAAGAATAGCGAAGGGCAAGAAGACAACGGCCCT

1501 AACTGGCGAAATTTGACTTTACCTGGTTTATTCCCTGCCATTATAAAATACAGGAGAATAT
-----+-----+-----+-----+-----+-----+ 1560
TTGACCGCTTTAAACTGAAATGGACCAAATAAGGACGGTAATATTTATGTCCTCTTATA

1561 TTATTGAAACCCCTTGTGTGTCTGTTTTTTTACAATTATTTCATTATAAACCCCCCTTT
-----+-----+-----+-----+-----+-----+ 1620
AATAACTTTGGGAACAACACAGACAAAAAATGTTAATAAACGTAATTATTGGGGGGAA

1621 TTTTTCAGGTGGTTATGGACAAAGTATTAGTGCACAGGGGATTTTCAACTCTTAATGTTA
-----+-----+-----+-----+-----+-----+ 1680
AAAAAGTCCACCAATACCTGTTTCATAATCACGTGTCCCCTAAAAGTTGAGAATTACAAT

1681 TTACTGTCGCATTATCTGTTGTGGTGGTGGTTGAGATTATACTCAGCGGTTTAAGAACTT
-----+-----+-----+-----+-----+-----+ 1740
AATGACAGCGTAATAGACAACACCACCACAACTCTAATATGAGTCGCCAAATTCGTGA

1741 ACATTTTTCACATAGTACAAGTCGGATTGATGTTGAGTTGGGTGCCAAACTCTTCCGGC
-----+-----+-----+-----+-----+-----+ 1800
TGTA AAAACGTGTATCATGTTTCAGCCTAACTACAACCTCAACCCACGGTTTGAGAAGGCCG

1801 ATTTACTGGCGCTACCGATCTCTTATTTTGAGAGTCGTCGTGTTGGTGATACTGTTGCCA
-----+-----+-----+-----+-----+-----+ 1860
TAAATGACCGCGATGGCTAGAGAATAAACTCTCAGCAGCACAACTATGACAACGGT

1861 GGGTAAGAGAATTAGACCAGATCCGTAATTTCTGACAGGACAGGCATTAAACATCTGTTC
-----+-----+-----+-----+-----+-----+ 1920
CCCATTCTCTTAATCTGGTCTAGGCATTAAAAGACTGTCCTGTCCGTAATTGTAGACAAG

HindIII
      |
1921 TGGACTTATTATTTTCATTCAATTTTGTGCGGTAATGTGGTATTACAGTCCAAAGCTT
-----+-----+-----+-----+-----+-----+ 1979
ACCTGAATAATAAAAGTAAGTATAAAAAACGCCATTACACCATAATGTCAGGTTTCGAA

```

Fig. 7 (cont.)

17/30

Enzimas que cortan:

BamHI	BglII	EagI	EcoRI	HindIII	NcoI	NheI	SalI
SfiI	XbaI	XmaI					

Enzimas que no cortan:

Ninguno

Fig. 7

18/30

Mapa de pVamyHLYA

Con 8 enzimas: NCOI PSTI SALI HINDIII SFII BAMHI NOTI ECORI PSTI

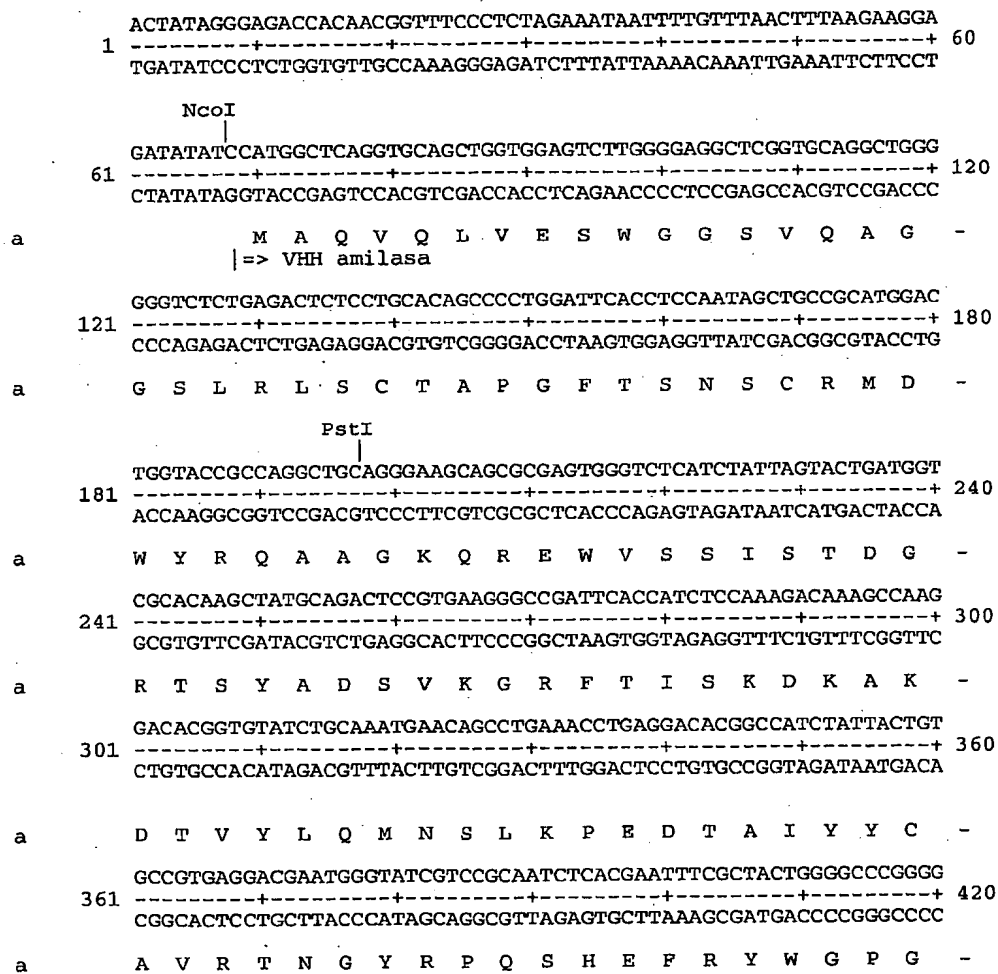


Fig. 8 (cont.)

SfiI
SalI

421 ACCCAGGTCACCGTCTCCTCAACGGCCTCGGGGGCCGCTCGACGCCCGGGGGTGCGCCG 480
 -----+-----+-----+-----+-----+-----+-----+-----+
 TGGGTCCAGTGGCAGAGGAGTTGCCGGAGCCCCCGGCGCAGCTGCGGGCCCCCACCGCGC
 a T Q V T V S S T A S G A A S T P G G A P -
BamHI
|=> E-tag

481 GTGCCGTATCCGGATCCGCTGGAACCGGCCGGGgAAAATtCTCTTGCTAAAAATGTATTA 540
 -----+-----+-----+-----+-----+-----+-----+
 CACGGCATAGGCCCTAGGCGACCTTGGCCGGCCCCtTTTaaGAGAACGATTTTTACATAAT
 a V P Y P D P L E P A G E N S L A K N V L -
|=> C-hlyA

541 TCCGGTGGAAAAGGTAATGACAAGTTGTACGGCAGTGAGGGAGCAGACCTGCTTGATGGC 600
 -----+-----+-----+-----+-----+-----+-----+
 AGGCCACCTTTTCCATTACTGTTCACATGCCGTCACCTCCCTCGTCTGGACGAACCTACCG
 a S G G K G N D K L Y G S E G A D L L D G -
 601 GGAGAAGGGAATGATCTTCTGAAAGGTGGATATGGTAATGATATTTATCGTTATCTTTCA 660
 -----+-----+-----+-----+-----+-----+-----+
 CCTCTTCCCTTACTAGAAGACTTTCACCTATACCATTACTATAAATAGCAATAGAAAGT
 a G E G N D L L K G G Y G N D I Y R Y L S -
 661 GGATATGGCCATCATATTATTGACGATGAAGGGGGGAAAGACGATAAACTAGTTTAGCT 720
 -----+-----+-----+-----+-----+-----+-----+
 CCTATACCGGTAGTATAATAACTGCTACTTCCCCCTTCTGCTATTTGAGTCAAATCGA
 a G Y G H H I I D D E G G K D D K L S L A -
 721 GATATAGATTTCCGGGACGTTGCCTTTAAGCGAGAAGGGAATGACCTCAITATGTATAAA 780
 -----+-----+-----+-----+-----+-----+-----+
 CTATATCTAAAGGCCCTGCAACGGAAATTCGCTCTTCCCTTACTGGAGTAATACATATTT
 a D I D F R D V A F K R E G N D L I M Y K -
 781 GCTGAAGGTAATGTTCTTCTTATGGCCACAAAAATGGTATTACATTTAAAAACTGGTTT 840
 -----+-----+-----+-----+-----+-----+-----+
 CGACTTCCATTACAAGAAAGATAACCGGTGTTTTTACCATAATGTAAATTTTTGACCAAA
 a A E G N V L S I G H K N G I T F K N W F -
 841 GAAAAAGAGTCAGATGATCTCTCTAATCATCAGATAGAGCAGATTTTGTATAAAGACGGC 900
 -----+-----+-----+-----+-----+-----+-----+
 CTTTTTCTCAGTCTACTAGAGAGATTAGTAGTCTATCTCGTCTAAAACTATTCTGCGC
 a E K E S D D L S N H Q I E Q I F D K D G -
 901 AGGGTAATCACACCAGATTCTCTTAAAAAAGCATTGTAATATCAGCAGAGTAATAACAAG 960
 -----+-----+-----+-----+-----+-----+-----+
 TCCCATTAGTGTTGCTCTAAGAGAATTTTTTCGTAAACTTATAGTCGTCTCATTTATGTTT

Fig. 8 (cont.)

20/30

a R V I T P D S L K K A F E Y Q Q S N N K -
GTAAGTTATGTGTATGGACATGATGCATCAACTTATGGGAGCCAGGACAATCTTAATCCA
961 -----+-----+ 1020
CATTCAATACACATACCTGTACTACGTAGTTGAATACCCTCGGTCTGTTAGAATTAGGT

a V S Y V Y G H D A S T Y G S Q D N L N P -
PstI
TTAATTAATGAAATCAGCAAAATCATTTCAGCTGCAGGTAACCTTCGATGTTAAGGAGGAA
1021 -----+-----+ 1080
AATTAATTACTTTAGTCGTTTGTAGTAAAGTCGACGTCCATTGAAGCTACAATTCCTCCTT

a L I N E I S K I I S A A G N F D V K E E -
AGATCTGCCGCTTCTTTATTGTCAGTTGTCCGGTAATGCCAGTGATTTTTCATATGGACGG
1081 -----+-----+ 1140
TCTAGACGGCGAAGAAATAACGTCAACAGGCCATTACGGTCACTAAAAAGTATACCTGCC

a R S A A S L L Q L S G N A S D F S Y G R -
AACTCAATAACTTTGACAGCATCAGCATAATATATTAATTTAAATGATAGCAATCTTACT
1141 -----+-----+ 1200
TTGAGTTATTGAAACTGTCGTAGTCGTATTATATAATTAATTTACTATCGTTAGAATGA

a N S I T L T A S A * -
GGGCTGTGCCACATAAGATTGCTATTTTTTTTGGAGTCATAATGGATTCTTGTCATAAAAT
1201 -----+-----+ 1260
CCCGACACGGTGTATTCTAACGATAAAAAACCTCAGTATTACCTAAGAACAGTATTTA

1261 TGATTATGGGTTATACGCCCTGGAGATTTTAGCCCAATACCATAACGTCCTGTTAACCC
-----+-----+ 1320
ACTAATACCCAATATGCGGGACCTCTAAATCGGGTTATGGTATTGCAGAGACAATTGGG

1321 GGAAGAAATTAACATAGATTTGACACAGACGGGACTGGTCTGGGATTAACGTCATGGTT
-----+-----+ 1380
CCTTCTTTAATTTGTATCTAAACTGTGTCTGCCCTGACCAGACCCTAATTGCAGTACCAA

1381 GCTTGCTGCGAAATCTTTAGAACTAAAGGTAAACAGGTAAAAAACAATTGACCGATT
-----+-----+ 1440
CGAACGACGCTTTAGAAATCTTGATTCCATTTTGTCCATTTTTTTTGTTAAGTGGCTAA

1441 AAACTTTATTTCTCTGCCCGCATTTAGTCTGGAGAGAGGATGGACGTCATTTTATTCTGAC
-----+-----+ 1500
TTTGAAATAAAGAGACGGGCGTAATCAGACCTCTCTCCTACCTGCAGTAAAAATAAGACTG

Fig. 8 (cont.)

21/30

1501 TAAAGTCAGTAAAGAAGCAAACAGATATCTTATTTCTGATCTGGAGCAGCGAAATCCCCG 1560
-----+-----+-----+-----+-----+
ATTTTCAGTCATTTCTTCTGTTTGTCTATAGAATAAAGACTAGACCTCGTCGCTTTAGGGGC

1561 TGTTCCTCGAACAGTCTGAGTTTGAGGCGTTATATCAGGGGCATATATTCTTATCGCTTC 1620
-----+-----+-----+-----+-----+
ACAAGAGCTTGTCTAGACTCAAACCTCCGCAATATAGTCCCCGTATAATAAGAATAGCGAAG

1621 CCGTTCTTCTGTTGCCGGGAAACTGGCGAAATTTGACTTTACCTGGTTTATTCCTGCCAT 1680
-----+-----+-----+-----+-----+
GGCAAGAAGACAACGGCCCTTTGACCGCTTTAAACTGAAATGGACCAAATAAGGACGGTA

1681 TATAAAATACAGGAGAATATTTATTGAAACCCCTTGTGTGTCTGTTTTTTTTACAATTATT 1740
-----+-----+-----+-----+-----+
ATATTTTATGTCTCTTATAAATAACTTTGGGAACAACACAGACAAAAAATGTTAATAA

1741 TGCATTAATAACCCCCCTTTTTTTTCAGGTGGTTATGGACAAAGTATTAGTGACAGGGG 1800
-----+-----+-----+-----+-----+
ACGTAATATTGGGGGGGAAAAAAGTCCACCAATACCTGTTTCATAATCACGTGTCCCC

1801 ATTTTCAACTCTTAATGTTATTACTGTGCGATTATCTGTTGTGGTGGTGGTTTGAGATTAT 1860
-----+-----+-----+-----+-----+
TAAAAGTTGAGAATTACAATAATGACAGCGTAATAGACAACACCACCACAACTCTAATA

1861 ACTCAGCGGTTTAAGAACTTACATTTTTGCACATAGTACAAGTCGGATTGATGTTGAGTT 1920
-----+-----+-----+-----+-----+
TGAGTCGCCAAATTCTTGAATGTAAAACGTGTATCATGTTTCAGCCTAACTACAACCTCAA

1921 GGGTGCCAAACTCTCCGGCATTTACTGGCGCTACCGATCTCTTATTTTGAGAGTCGTCG 1980
-----+-----+-----+-----+-----+
CCCACGGTTTGAGAAGGCCGTAAATGACCGCGATGGCTAGAGAATAAACTCTCAGCAGC

1981 TGTGGTGATACTGTTGCCAGGGTAAGAGAATTAGACCAGATCCGTAATTTTCTGACAGG 2040
-----+-----+-----+-----+-----+
ACAACCACATGACAACGGTCCCATTCTCTTAATCTGGTCTAGGCATTAAAAGACTGTCC

2041 ACAGGCATTAAACATCTGTTCTGGACTTATTATTTTCATTCATATTTTTCGGGTAATGTG 2100
-----+-----+-----+-----+-----+
TGTCCGTAATTGTAGACAAGACCTGAATAATAAAAGTAAGTATAAAAAACGCCATTACAC

Fig. 8 (cont.)

22/30

HindIII
|

2101 GTATTACAGTCCAAAGCTTACTCTGGTGATCTTATTTTCGCTGCCTTGTATGCTGCATG 2160
 -----+-----+-----+-----+-----+-----+-----+
 CATAATGT CAGGTTTCGAATGAGACCAC TAGAATAAAAGCGACGGAACAATACGACGTAC

2161 GTCTGTTTTTATTAGCCCCATTTTGGCGACGTCGCCTTGATGATAAGTTTTCACGGAATGC 2220
 -----+-----+-----+-----+-----+-----+-----+
 CAGACAAAATAATCGGGGTAAAACGCTGCAGCGGAAC TACTATTCAAAGTGCCTTACG

2221 GGATAATCAATCTTTCCTGGTGAATCAGTCACGGCGATTAACTATAAAAGCTATGGC 2280
 -----+-----+-----+-----+-----+-----+-----+
 CCTATTAGTTAGAAAGGACCACCTTAGTCAGTGCCTAATTGTGATATTTTCGATACCG

PstI
|

2281 AGTCTCACCTCAGATGACGAACATATGGGACAAACAATTGGCAGGATATGTTGCTGCAGG 2340
 -----+-----+-----+-----+-----+-----+-----+
 TCAGAGTGGAGTCTACTGCTTGTATACCCTGTTTGTTAACCGTCCTATACAACGACGTCC

2341 CTTCAAAGTGACAGTATTAGCAACCATTGGTCAACAAGGAATACAGTTAATACAAAAGAC 2400
 -----+-----+-----+-----+-----+-----+-----+
 GAAGTTTCACTGTCTAATCGTTGGTAACCAAGTTGTTCCCTTATGTCAATTATGTTTCTG

2401 TGTATGATCATCAACCTGTGGTTGGGAGCACACCTGGTTATTTCCGGGGATTTAAGTAT 2460
 -----+-----+-----+-----+-----+-----+-----+
 ACAATACTAGTAGTTGGACACCAACCTCGTGTGGACCAATAAAGGCCCTAAATTCATA

2461 TGGTCAGTTAATTGCTTTTAATATGCTTGCTGGTCAGATTGTTGCACCGGTTATTCGCCCT 2520
 -----+-----+-----+-----+-----+-----+-----+
 ACCAGTCAATTAACGAAAATTATACGAACGACCACTCTAACAACGTGGCCAATAAGCGGA

2521 TGCACAAATCTGGCAGGATTTCCAGCAGGTTGGTATATCAGTTACCCGCCTTGGTGATGT 2580
 -----+-----+-----+-----+-----+-----+-----+
 ACGTGTTTAGACCGTCCTAAAGGTCGTCCAACCATATAGTCAATGGGCGGAACCACTACA

2581 GCTTAACCTCTCCAAC TGAAAGTTATCATGGGAAACTGGCATTACCGGAAATTAATGGTGA 2640
 -----+-----+-----+-----+-----+-----+-----+
 CGAATTGAGAGGTTGACTTTCAATAGTACCCTTTGACCGTAATGGCCTTTAATTACCCT

2641 TATCACTTTTCGTAATATCCGGTTTCGCTATAAGCCTGACTCTCCGGTTATTTTAGATAA 2700
 -----+-----+-----+-----+-----+-----+-----+
 ATAGTGAAAAGCATTTATAGGCCAAAGCGATATTTCGGACTGAGAGGCCAATAAAATCTATT

2701 TATCAATCTCAGTATTAAGCAGGGGGAGGTTATTGGTATTGTCGGACGTTCTGGTTCAGG 2760
 -----+-----+-----+-----+-----+-----+-----+
 ATAGTTAGAGTCATAATTCGTCCCCCTCCAATAACCATAACAGCCTGCAAGACCAAGTCC

Fig. 8 (cont.)

23/30

2761 AAAAAGCACATTAAGTAAATTAATTCAACGTT 2792
-----+-----+-----+---
TTTTTCGTGTAATTGATTTAATTAAAGTTGCAA

Enzimas que cortan:

BamHI HindIII NcoI PstI Sali SfiI

Enzimas que no cortan:

EcoRI NotI

Fig. 8

24/30

Mapa de pVamvZHLVA

Con 11 enzimas: ECORI SPHI PSTI NCOI NHEI NDEI BAMHI HINDIII
SALI SFII NOTI

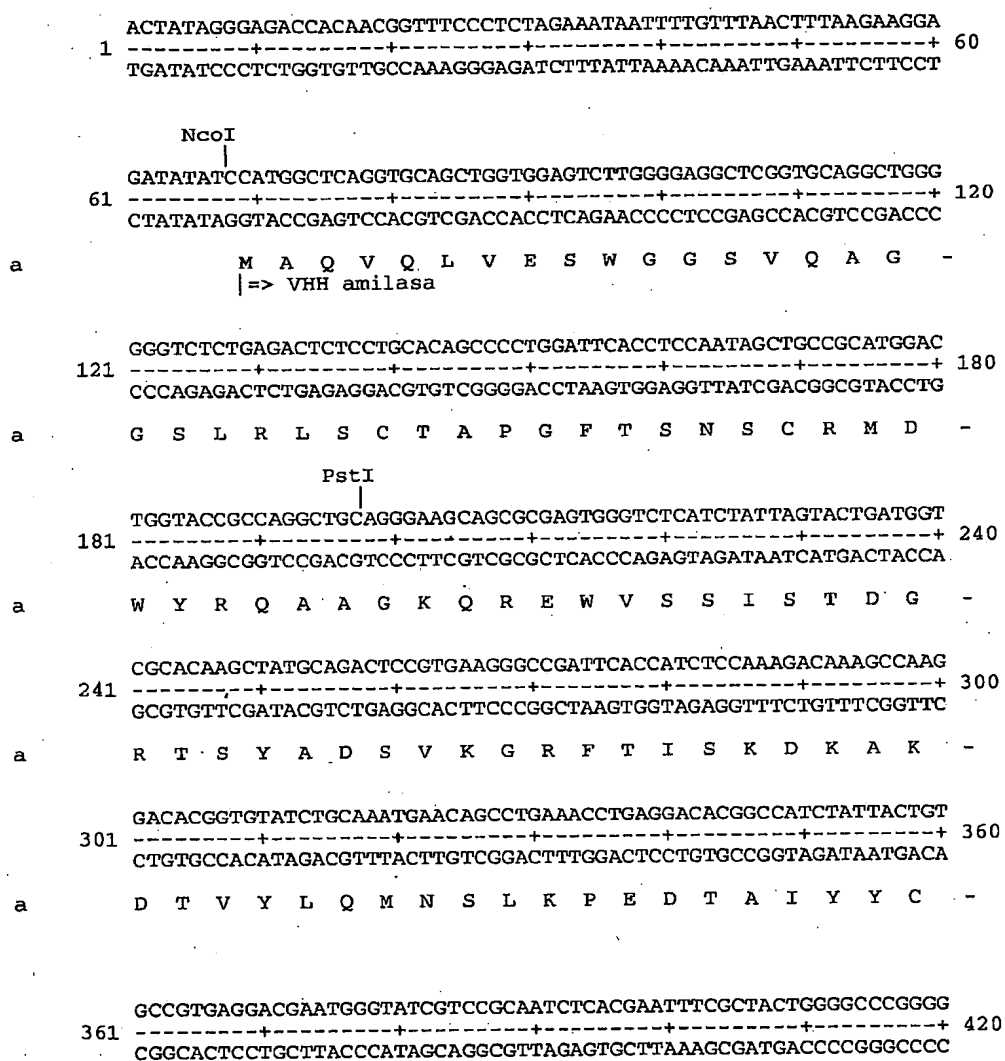


Fig. 9 (cont.)

25/30

```

a      A V R T N G Y R P Q S H E F R Y W G P G -

                                     sfII  SalI
                                     |    |
421  ACCCAGGTCACCGTCTCCTCAACGGCCTCGGGGGCCGCGTCGACGTCCGGCGGTCCGAAG
-----+-----+-----+-----+-----+-----+-----+-----+ 480
TGGGTCCAGTGGCAGAGGAGTTGCCGGAGCCCCCGGCGCAGCTGCAGGCCGCCAGGCTTC

a      T Q V T V S S T A S G A A S T S G G P K -
                                     |=> Ig hinge

481  CCTTCCACTCCGCCCGGGTCTTCCCGTATGAAACAGCTGGAAGACAAAGTAGAGGAGCTC
-----+-----+-----+-----+-----+-----+-----+ 540
GGAAGGTGAGGCGGGCCCCAGAAGGGCATACTTTGTCGACCTTCTGTTCATCTTCTCGAG

a      P S T P P G S S R M K Q L E D K V E E L -
                                     |=> Leucine Zipper GCN4

                                     HindIII
                                     |
541  CTTAGCAAGAACTACCATCTAGAAAACGAGGTAGCTCGTCTGAAAAAGCTTGTGGTGAA
-----+-----+-----+-----+-----+-----+-----+ 600
GAATCGTTCTTGATGGTAGATCTTTTGCTCCATCGAGCAGACTTTTTCGAACAACCACITT

a      L S K N Y H L E N E V A R L K K L V G E -

                                     SalI
                                     |
601  CGTGGTGGTCACCATCACCATCACCATGCGTCGACGCCCGGGGGTGCGCCGGTGCCGTAT
-----+-----+-----+-----+-----+-----+-----+ 660
GCACCACCAGTGGTAGTGGTAGTGGTACGCAGCTGCGGGCCCCCACGCGGCCACGGCATA

a      R G G H H H H H H A S T P G G A P V P Y -
      |=> 6xhis tag                                     |=> E-tag

BamHI
|
661  CCGGATCCGCTGGAACCGGCCGGGgAAAATtCTCTTGCTAAAAATGTATTATCCGGTGGA
-----+-----+-----+-----+-----+-----+-----+ 720
GGCCTAGGCGACCTTGGCCGGCCCCtTTTAaGAGAACGATTTTACATAATAGGCCACCT

a      P D P L E P A G E N S L A K N V L S G G -
                                     |=> C-hlyA

721  AAAGGTAATGACAAGTTGTACGGCAGTGAGGGAGCAGACCTGCTTGATGGCGGAGAAGGG
-----+-----+-----+-----+-----+-----+-----+ 780
TTTCCATTACTGTTCAACATGCCGTCACCTCCCTCGTCTGGACGAACCTACCGCCTCTTCCC

a      K G N D K L Y G S E G A D L L D G G E G -

```

Fig. 9 (cont.)

26/30

781 AATGATCTTCTGAAAGGTGGATATGGTAATGATATTTATCGTTATCTTTCAGGATATGGC 840
-----+-----+-----+-----+-----+
TTACTAGAAGACTTTCCACCTATACCATTACTATAAATAGCAATAGAAAGTCCTATACCG
a N D L L K G G Y G N D I Y R Y L S G Y G -

841 CATCATATTATTGACGATGAAGGGGGGAAAGACGATAAACTCAGTTTAGCTGATATAGAT 900
-----+-----+-----+-----+-----+
GTAGTATAATAACTGCTACTTCCCCCTTTCTGCTATTTGAGTCAAATCGACTATATCTA
a H H I I D D E G G K D D K L S L A D I D -

901 TTCCGGGACGTTGCCTTTAAGCGAGAAGGGAATGACCTCATTATGTATAAAGCTGAAGGT 960
-----+-----+-----+-----+-----+
AAGGCCCTGCAACGGAATTTCGCTCTTCCCTTACTGGAGTAATACATATTTGCACTTCCA
a F R D V A F K R E G N D L I M Y K A E G -

961 AATGTTCTTTCTATTTGGCCACAAAAATGGTATTACATTTAAAACTGGTTTGAAAAAGAG 1020
-----+-----+-----+-----+-----+
TTACAAGAAAGATAACCGTGTTTTTACCATAATGTAAATTTTGACCAAACCTTTTTCTC
a N V L S I G H K N G I T F K N W F E K E -

1021 TCAGATGATCTCTCTAATCATCAGATAGAGCAGATTTTGGATAAAGACGGCAGGGTAATC 1080
-----+-----+-----+-----+-----+
AGCTACTAGAGAGATTAGTAGTCTATCTCGTCTAAAACTATTTCTGCCGTCCCATTAG
a S D D L S N H Q I E Q I F D K D G R V I -

1081 ACACCAGATTCTCTTAAAAAAGCATTGGAATATCAGCAGAGTAATAACAAGGTAAGTTAT 1140
-----+-----+-----+-----+-----+
TGTGGTCTAAGAGAATTTTTCGTAAACTTATAGTCGTCTCATTATGTTCCATTCAATA
a T P D S L K K A F E Y Q Q S N N K V S Y -

1141 GTGTATGGACATGATGCATCAACTTATGGGAGCCAGGACAATCTTAATCCATTAATTAAT 1200
-----+-----+-----+-----+-----+
CACATACCTGTACTACGTAGTTGAATACCCTCGGTCTGTAGAAATTAGGTAATTAATTA
a V Y G H D A S T Y G S Q D N L N P L I N -

Fig. 9 (cont.)

27/30

PstI

1201 GAAATCAGCAAAATCATTTCAGCTGCAGGTAACCTTCGATGTTAAGGAGGAAAGATCTGCC 1260
 -----+-----+-----+-----+-----+-----+-----+
 CTTTAGTCGTTT TAGTAAAGTCGACGTCCATTGAAGCTACAATTCCTCCTTTCTAGACGG

a E I S K I I S A A G N F D V K E E R S A -

NdeI

1261 GCTTCTTTATTCAGTTGTCCGGTAATGCCAGTGATTTTTCATATGGACGGAACCAATA 1320
 -----+-----+-----+-----+-----+-----+-----+
 CGAAGAAATAACGTCAACAGGCCATTACGGTCACTAAAAAGTATACCTGCCTTGAGTTAT

a A S L L Q L S G N A S D F S Y G R N S I -

1321 ACTTTGACAGCATCAGCATAATATATTAATTTAAATGATAGCAATCTTACTGGGCTGTGC 1380
 -----+-----+-----+-----+-----+-----+-----+
 TGAAACTGTCGTAGTCGTATTATATAATTAAATTTACTATCGTTAGAATGACCCGACACG

a T L T A S A * -

1381 CACATAAGATTGCTATTTTTTGGAGTCATAATGGATTCTTGTCATAAAATTGATTATGG 1440
 -----+-----+-----+-----+-----+-----+-----+
 GTGTATTCTAACGATAAAAAAACCTCAGTATTACCTAAGAACAGTATTTTAACATAATACC

1441 GTTATACGCCCTGGAGATTTTAGCCCAATACCATAACGTCTCTGTTAACCCGGAAGAAAT 1500
 -----+-----+-----+-----+-----+-----+-----+
 CAATATGCGGGACCTCTAAAATCGGGTTATGGTATTGCAGAGACAATTGGGCCCTCTTTA

1501 TAAACATAGATTTGACACAGACGGGACTGGTCTGGGATTAACGTCATGGTTGCTTGCTGC 1560
 -----+-----+-----+-----+-----+-----+-----+
 ATTTGTATCTAAACTGTGTCTGCCCTGACCAGACCCTAATTGCAGTACCAACGAACGACG

1561 GAAATCTTTAGAACTAAAGGTAAAAACAGGTAAAAAAAACAATTGACCGATTAACTTTAT 1620
 -----+-----+-----+-----+-----+-----+-----+
 CTTTAGAAATCTTGATTTCATTTTGTCCATTTTGTGTTAACTGGCTAATTTGAAATA

1621 TTCTCTGCCCGCATTAGTCTGGAGAGAGGATGGACGTCATTTTATTCTGACTAAAGTCAG 1680
 -----+-----+-----+-----+-----+-----+-----+
 AAGAGACGGGCGTAATCAGACCTCTCTCTACCTGCAGTAAATAAGACTGATTTTCAGTC

1681 TAAAGAAGCAAACAGATATCTTATTTCTGATCTGGAGCAGCGAAATCCCCGTGTTCTCGA 1740
 -----+-----+-----+-----+-----+-----+-----+
 ATTTCTTCGTTTGTCTATAGAATAAAGACTAGACCTCGTCGCTTTAGGGGCACAAGAGCT

Fig. 9 (cont.)

28/30

1741 ACAGTCTGAGTTTGGAGCGTTATATCAGGGGCATATTATTCTTATCGCTTCCCGTTCTTC 1800
-----+-----+-----+-----+-----+-----+
TGTCAGACTCAAACCTCCGCAATATAGTCCCGTATAATAAGAATAGCGAAGGGCAGAAG

1801 TGTGTGCCGGGAAACTGGCGAAATTTGACTTTACCTGGTTTATTCCTGCCATTATAAAATA 1860
-----+-----+-----+-----+-----+-----+
ACAACGGCCCTTTGACCGCTTTAAACTGAAATGGACCAAATAAGGACGGTAATATTTTAT

1861 CAGGAGAATATTTATTGAAACCCTGTTGTGTCGTGTTTTTTTACAATTTATTTGCATTAAT 1920
-----+-----+-----+-----+-----+-----+
GTCCTCTTATAAATAACTTTGGGAACAACACAGACAAAAAATGTTAATAAACGTAATTA

1921 AACCCCCCTTTTTTTCAGGTGGTTATGGACAAAGTATTAGTGCACAGGGGATTTTCAAC 1980
-----+-----+-----+-----+-----+-----+
TTGGGGGGAAAAAAGTCCACCAATACCTGTTTCATAATCAGTGTCCCTAAAGTTG

1981 TCTTAATGTTATTACTGTGCGATTATCTGTTGTGGTGGTGTGTTGAGATTATACTCAGCGG 2040
-----+-----+-----+-----+-----+-----+
AGAATTACAATAATGACAGCGTAATAGACAACACCACCACAAACTCTAATATGAGTCGCC

2041 TTTAAGAAGTTACATTTTGCACATAGTACAAGTCGGATTGATGTTGAGTTGGGTGCCAA 2100
-----+-----+-----+-----+-----+-----+
AAATTCCTGAATGTA AAAACGTGTATCATGTT CAGCCTAACTACAAC TCAACCCACGGTT

2101 ACTCTTCCGGCATTACTTGGCGCTACCGATCTCTTATTTTGGAGAGTCGTCGTGTTGGTGA 2160
-----+-----+-----+-----+-----+-----+
TGAGAAGGCCGTAAATGACCGCGATGGCTAGAGAAATAAACTCTCAGCAGCACAAACCACT

2161 TACTGTTGCCAGGGTAAGAGAATTAGACCAGATCCGTAATTTTCTGACAGGACAGGCAIT 2220
-----+-----+-----+-----+-----+-----+
ATGACAACGGTCCCATTCTCTTAATCTGGTCTAGGCATTAAAAGACTGTCCTGTCCGTAA

2221 AACATCTGTTCTGGACTTATTATTTTCATTCATATTTTTTGGCGTAATGTGGTATTACAG 2280
-----+-----+-----+-----+-----+-----+
TTGTAGACAAGACCTGAATAATAAAAGTAAGTATAAAAAACGCCATTACACCATAATGTC

HindIII
|
2281 TCCAAAGCTTACTCTGGTGATCTTATTTTCGCTGCCTTGTTATGCTGCATGGTCTGTTTT 2340
-----+-----+-----+-----+-----+-----+
AGGTTTCGAATGAGACCACTAGAATAAAAGCGACGGAACAATACGACGTACCAGACAAAA

2341 TATTAGCCCCATTTTGGCAGCTCGCCTTGATGATAAGTTTTCACGGAATGCGGATAATCA 2400
-----+-----+-----+-----+-----+-----+
ATAATCGGGGTAAAACGCTGCAGCGGAAC TACTATTCAAAGTGCCTTACGCCTATTAGT

Fig. 9 (cont.)

29/30

```
2401 ATCTTTCCTGGTGGAATCAGTCACGGCGATTAACTATAAAAGCTATGGCAGTCTCACC 2460
      -----+-----+-----+-----+-----+-----+
      TAGAAAGGACCACCTTAGTCAGTGCCGCTAATTGTGATATTTTCGATACCGTCAGAGTGG

                NdeI                                PstI
2461 TCAGATGACGAACATATGGGACAAACAATTGGCAGGATATGTTGCTGCAGGCTTCAAAGT 2520
      -----+-----+-----+-----+-----+-----+
      AGTCTACTGCTTGTATACCCCTGTTTGTTAACCGTCCTATACAACGACGTCCGAAGTTTCA

2521 GACAGTATTAGCAACCATTTGGTCAACAAGGAATACAGTTAATACAAAAGACTGTTATGAT 2580
      -----+-----+-----+-----+-----+-----+
      CTGTCATAATCGTTGGTAACCAGTTGTTTCTTATGTCAATTATGTTTCTGACAATACTA

2581 CATCAACCTGTGGTTGGGAGCACACCTGGTTATTTCCGGGGATTAAAGTATTGGTCAGTT 2640
      -----+-----+-----+-----+-----+-----+
      GTAGTTGGACACCAACCCTCGTGTGGACCAATAAAGGCCCTAAATTCATAACCAGTCAA

2641 AATTGCTTTTAAATATGCTTGGTCAGATTGTTGCACCGGTTATTCGCCTTGCACAAAT 2700
      -----+-----+-----+-----+-----+-----+
      TTAACGAAAATTATACGAACGACCAGTCTAACAACGTGGCCAATAAGCGGAACGTGTTTA

2701 CTGGCAGGATTTCCAGCAGGTTGGTATATCAGTTACCCGCCTTGGTGATGTGCTTAACTC 2760
      -----+-----+-----+-----+-----+-----+
      GACCGTCCTAAAGGTCGTCCAACCATATAGTCAATGGGCGGAACCACACACGAATTGAG

2761 TCCAACTGAAAGTTATCATGGGAACTGGCATTACCGGAAATTAATGGTGATATCACTTT 2820
      -----+-----+-----+-----+-----+-----+
      AGGTTGACTTTCAATAGTACCCCTTGACCGTAATGGCCTTTAATTACCACTATAGTAAA

2821 TCGTAATATCCGGTTTCGCTATAAGCCTGACTCTCCGGTTATTTTAGATAATATCAATCT 2880
      -----+-----+-----+-----+-----+-----+
      AGCATTATAGGCCAAAGCGATATTCGACTGAGAGGCCAATAAAATCTATTATAGTTAGA

2881 CAGTATTAAGCAGGGGAGGTTATTGGTATTTGTCGGACGTTCTGGTTCAGGAAAAAGCAC 2940
      -----+-----+-----+-----+-----+-----+
      GTCATAATTCGTCCCCCTCCAATAACCATAACAGCCTGCAAGACCAAGTCCTTTTTCGTG

2941 ATTAACATAATTAATCAACGTT 2963
      -----+-----+-----
      TAATTGATTTAATTAAGTTGCAA
```

Fig. 9 (cont.)

30/30

Enzimas que cortan:

BamHI HindIII NcoI NdeI PstI SalI SfiI

Enzimas que no cortan:

EcoRI NheI NotI SphI

Fig. 9